ACE 592 ME: Lecture Notes

Barrett Kirwan

August 26, 2015

8b42da3 on 2015/08/26

Contents

Part I

Preliminaries

1 Proving Conditional Statements

Book of Proof, chpts. 4-6

> The trouble with college math classes—which classes consist almost entirely in the rhythmic ingestion and regurgitation of abstract information, and are paced in such a way as to maximize this reciprocal dataflow—is that their sheer surface-level difficulty can fool us into thinking we really know something when all we really 'know' is abstract formulas and rules for their deployment. Rarely do math classes ever tell us whether a certain formula is truly significant, or why, or where it came from, or what was at stake.[1] —David Foster Wallace [2]

A formal system of proof requires *axioms* and *rules of inference*. Axioms are basic propositions so obvious that they can be asserted without proof. Rules of inference, which are sometimes called the Laws of Thought, are the logical principles that justify deriving truths from other truths.

Rules of Inference

Law of Identity

If anything is P, then it is P.

Law of the Excluded Middle

By LEM[3], a mathematical proposition P must be either true or, if not true, false.[4]

[1] And, of course, rarely do students think to ask—the formulas alone take so much work to 'understand' (i.e., to be able to solve problems correctly with), we often aren't aware that we don't understand them at all. That we end up not even knowing that we don't know is the really insidious part of most math classes.

[2] For the most perspicuous treatment of calculus and the history of math, read Wallace (2003).

[3] This is one of the Three Classic Laws of Thought. Look it up on Wikipedia.

[4] The 'or if not true,' part is required in formal logic because of certain properties of the disjunctive operator 'or'.

Entailment

'If …then' and often represented by the symbol '$\rightarrow$'. The most obvious rule of entailment is that (1) $P \rightarrow Q$ and (2) P *is true* license the conclusion (3) Q *is true*. Keep in mind that entailment, as a logical relation, has to do not with cause but with surety. *Necessary*- and *sufficient condition* are the terms in applied logic. If, for example, P is taken to mean 'is 5 feet tall' and Q to mean 'is at least 4 feet 11 inches tall,' then the purely logical meaning of '$P \rightarrow Q$' becomes evident: it really means 'If P is true then there is no way Q can be false'.

Direct Proof

Proposition

If P, then Q.

P	Q	$P \rightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

Advice It's generally a good idea to write the first and last line of your proof first.

In mapping out a strategy, it can be helpful to work backwards.

Using cases

"Without loss of generality …" is a common way of signalling that the proof is treating just one of several nearly identical cases.

Contrapositive Proof

Modus tollens Latin for 'method of denying'. E.g., $P \rightarrow Q$ and Q *is false* license P *is false*.

Proposition

If P, then Q.

P	Q	$\sim Q$	$\sim P$	$\sim Q \to \sim P$	$P \to Q$
T	T	F	F	T	T
T	F	T	F	F	F
F	T	F	T	T	T
F	F	T	T	T	T

The expression $\sim Q \to \sim P$ is the *contrapositive* of $P \to Q$.[5]

Proof by Contradiction

These two rules, LEM and modus tollens, enable the method of *proof by contradiction*, also know as proof by *reductio ad absurdum* or simply *reductio*. Here's how it works. Say you want to prove Q. What you do is assume not-Q and then show that not-Q logically entails a contradiction like, say, 'P and not-P'. By modus tollens, if not-$Q \to (P$ & not-$P)$ and $(P$ & not-$P)$ is false (which it is by LEM), then not-Q is false; and, by LEM, if not-Q is false, then Q must be true.

Read through that again to let it sink in.

Exercise 1.1 Construct the "truth table" to show $(P \wedge \sim Q) \to (P \wedge \sim P)$.

Proposition

If P, then Q.

The first step is to assume P and $\sim Q$. (Don't assume $\sim P$ because then the proposition is trivially true (see Direct Proof truth table).)

Advice Despite the power of the *reductio*, it's best to use it only when the direct and contrapositive approaches do not seem to work. The reason for this is that a proof by contradiction can often have hidden in it a simpler contrapositive proof, and if this is the case, it's better to go with the simpler approach.

[5]Don't confuse the words *contrapositive* and *converse*. The *converse* of $P \to Q$ is the statement $Q \to P$, which is not logically equivalent to $P \to Q$.

Mathematical Induction

Mathematical induction is a proof method that proves two things:

1. The proposition holds for the smallest number
2. If the proposition is true for k then it is true for $k + 1$.

Note Note a couple of things about this technique.

- First, the phrase "if the proposition is true for k" is called the inductive hypothesis.
- Part 2 is all about proving the process from k to $k + 1$.

Proof by induction only works if there is a smallest number. For example, there is no smallest number in the set $(0, 1]$. So we need to know when to expect a smallest element.

Figure 1.1: Proof by intimidation (xkcd.com)

Part II

Producer Theory—A Traditional Approach

2 Production

**Jehle and Reny 3*
Rubinstein Lecture 7
MWG 5.B

The Firm's "Behavior"

Think about the firm as an economic agent whose primary objective is to make money (profit). But it isn't, really, despite what the Supreme Court has ruled and what Stephen Colbert claims. Nevertheless, in keeping with tradition, we are going to treat it as an economic agent.

A Mirror of Consumer Theory

The producer's problem is similar to the consumer's *dual* problem. In the theory of the firm, preferences are linear and the constraint is a convex set. In consumer theory the constraint is a linear inequality and preferences were convex. We impose structure (continuity and convexity) on the producer's choice set and on the consumer's preferences.

Alternative Behavior

Rubinstein (2012) provides several examples of alternative objectives a firm might have: (see p. 93)

Production Sets, I.E., *Technology*

Notation *Production vector / output vector / production plan* is $\mathrm{y} = (y_1, \ldots, y_L) \in \mathbb{R}^L$, and the set of all feasible production plans is $Y \subset \mathbb{R}^L$.

Production set is Y. If y is possible then we write $\mathrm{y} \in Y$. $\mathrm{y} \notin Y$ means y is not possible. The shape of Y is going to be driven by the way in which different inputs can be substituted for each other in the production process.

Often, we will distinguish between *inputs* and *outputs*. We will denote the production levels of the firm's M outputs as $\mathrm{q} = (q_1, \ldots, q_m) \geq 0$. The use level of the firm's $L - M$ inputs is denoted as $\mathrm{z} = (z_1, \ldots, z_{L-M}) \geq 0$. There are $M + (L - M) = L$ total types of goods.

If $M = 1$, we have a single output technology, often described by a *production function* q $= f(\mathrm{z})$. Thus, we can write

$$Y = \{(-z_1, \ldots, -z_{L-1}, q) | q - f(z_1, \ldots, z_{L-1}) \leq 0, (z_1, \ldots, z_{L-1}) \geq 0\}. \quad (2.1)$$

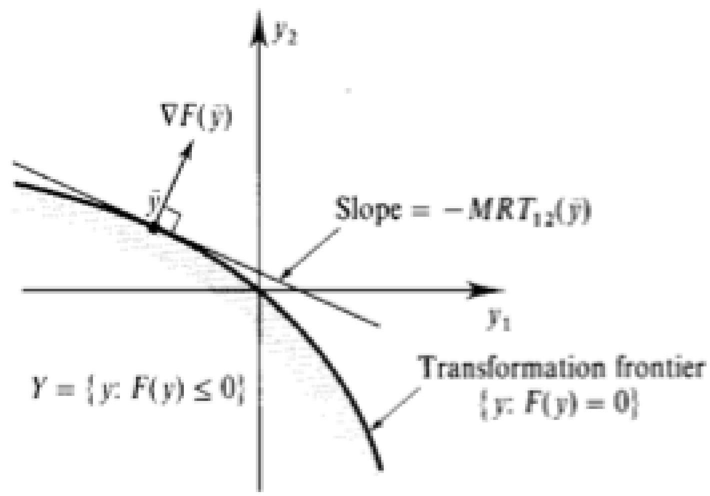

Figure 2.1: Transformation Frontier

Properties of Production Sets

Properties of all production sets

What follows are properties of a well-defined production set; often called the regularity conditions for this particular type of problem.

Nonempty $Y \neq \{\}$. Otherwise there is nothing to analyze.

Closed Y is closed, i.e., it contains its own boundary. If not, you may not be able to maximize a function subject to the constraint that the production plan be in Y—the firm will try to be as close to the boundary as possible, but no matter how close it is, it could always be a little closer.

No free lunch $f(0) \ngtr 0$.

The firm cannot produce output without using any inputs. Any feasible production plan y must have at least one negative component.

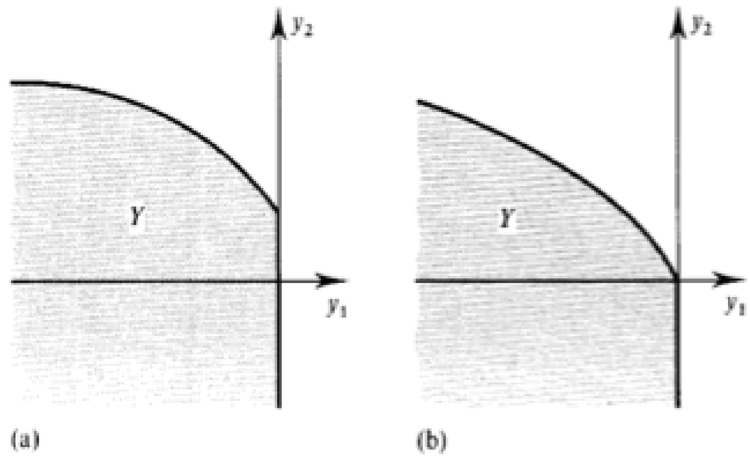

Figure 2.2: No Free Lunch

Free disposal For any point in Y, points that use less of all components are also in Y. It is always possible to just throw away inputs or outputs without using any inputs to do so. The production set is unbounded as you move down and to the left. This allows us to assume that a point that satisfies the first-order conditions is really a maximum.

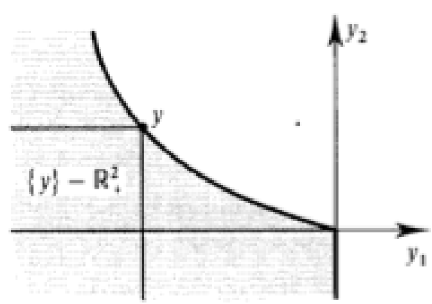

Figure 2.3: Free Disposal

Properties of some production sets

What follows are possible properties of a production set. Not all can apply at the same time. The appropriate properties will depend on the circumstances.

Irreversibility $y \in Y$ implies $-y \notin Y$, i.e., the production process cannot be undone. For example, you can use N fertilizer to make corn, but once N fert is applied, you can't simply reverse the process and give up corn to get back your N fert.

Exercise 1

Exercise 1. *Draw two production sets: one that violates irreversibility and one that satisfies this property.*

Inaction $0 \in Y$, i.e., there is a possibility to do nothing.

Global Returns to Scale Properties

No IRTS *Nonincreasing returns to scale.* $y \in Y$ and $\alpha \in [0, 1]$ implies $\alpha y \in Y$.

A technology that exhibits increasing returns to scale is one that becomes more productive as the size of the output grows. To rule this out, we require any feasible production plan y to be scaled down to ay for $a \in [0.1]$.

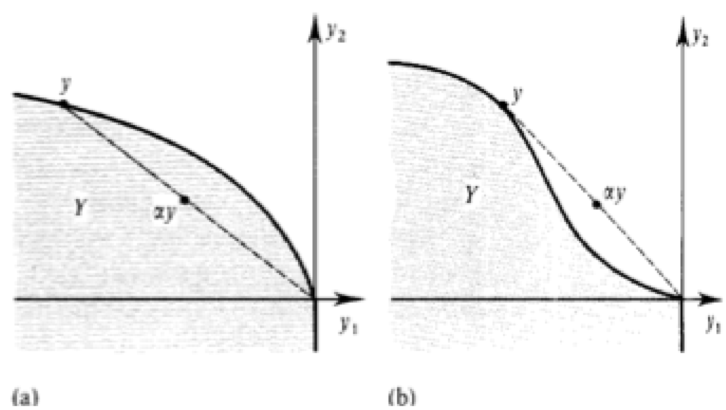

Figure 2.4: NonIncreasing Returns to Scale

No DRTS *Nondecreasing returns to scale.* $\mathrm{y} \in Y$ and $\alpha \geq 1$ implies $\alpha \mathrm{y} \in Y$.

Decreasing returns to scale is a situation where the firm grows less productive at higher levels of output. If a firm has fixed costs, it may exhibit nondecreasing returns to scale but cannot exhibit nonincreasing returns to scale.

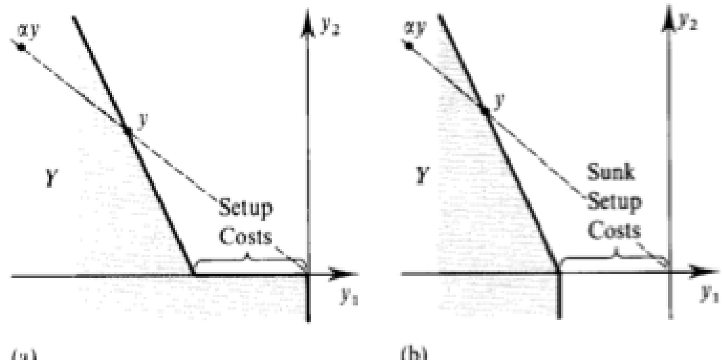

Figure 2.5: NonDecreasing Returns to Scale

CRTS *Constant returns to scale.* $\mathrm{y} \in Y$ and $\alpha \geq 0$ implies $\alpha \mathrm{y} \in Y$. RTS must be both nonincreasing and nondecreasing. The firm's productivity is independent of the level of production.

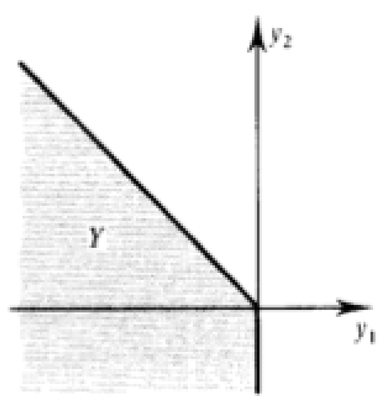

Figure 2.6: Constant Returns to Scale

Does production have to be linear for CRTS? Not if there is more than one input.

NB This technology has *both* CRTS and diminishing marginal returns.

Exercise 2 Suppose that $f(\cdot)$ is the production function associated with a single-output technology and let Y be the production set of this technology. Show that Y satisfies CRTS iff $f(\cdot)$ is homogeneous of degree one.

Free entry *Additivity*. $\mathrm{y} \in Y$ and $\mathrm{y}' \in Y$ implies $\mathrm{y} + \mathrm{y}' \in Y$. This implies $k\mathrm{y} \in Y$ for any positive integer k.

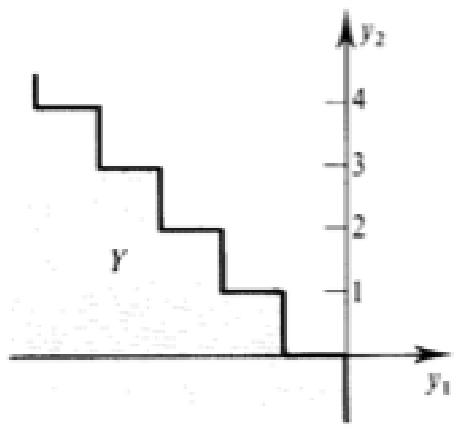

Figure 2.7: Additivity

Convexity Convexity is one of the *fundamental assumptions of microeconomics*. If Y exhibits nonincreasing returns to scale, then Y is convex.

$\mathrm{y} \in Y$,$\mathrm{y}' \in Y$ and $\alpha \in [0, 1]$ implies $\alpha\mathrm{y} + (1 - \alpha)\mathrm{y}' \in Y$.

This incorporates two things.

- nonincreasing RTS (if inaction is possible). As long as $0 \in Y$, then convexity implies $\alpha\mathrm{y} + (1 - \alpha)0 \in Y$ for all $\alpha \in [0, 1]$. So $\alpha\mathrm{y} \in Y$ for all $\alpha \in [0, 1]$, which is just the definition of nonincreasing RTS
- Balanced production plans produce more than unbalanced ones do.

Convex cone A convex cone is a conjunction of the convexity and CRTS properties: $\mathrm{y}, \mathrm{y}' \in Y$ and $\alpha, \beta \geq 0$ implies $\alpha\mathrm{y} + \beta\mathrm{y}' \in Y$.

Proposition (MWG 5.B.1) The production set Y is additive and satisfies the nonincreasing returns condition iff it is a convex cone.

Exercise 3 Prove the proposition.

Implications Remember that the production set describes *technology*, not limits on resources. If the production set is a convex cone, it implies that a farmer producing 1 million bushels of corn from $(\bar{x}_1, \ldots, \bar{x}_n)$ (inputs) should be able to do it again on another farm. If he doesn't, there must be some kind of input constraint $\rightarrow$ land?, entrepreneurship?

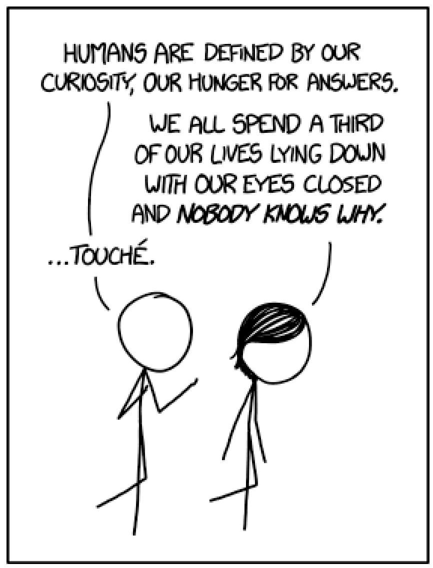

Figure 2.8: Answers (xkcd.com)

3 Concavity

**Jehle and Reny A1.4.2 and A1.4.3*
MWG M.C

(Quasi)Concave Functions, Convex Combinations, and Convex Sets

Before discussing some of the properties of production functions, let's be sure we're clear on some crucial definitions. These concepts are addressed in more detail on the following website:

http://www.economics.utoronto.ca/osborne/MathTutorial/QCC.HTM.

Concave Functions

Concave Function

$f : D \to \mathbb{R}$ is a *concave function* if for all $\mathrm{x}, \mathrm{y} \in D$,

$$f\left(t\mathrm{x} + (1-t)\mathrm{y}\right) \geq tf(\mathrm{x}) + (1-t)f(\mathrm{y}) \qquad \forall t \in [0,1] \tag{3.1}$$

Strict concavity occurs when the inequalities are strict.

Examples

A function is concave if every straight line connecting two points on the surface lies everywhere on or under the surface.

The above condition can be shown to be equivalent to

$$f\left(\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n\right) \geq \lambda_1 f(x_1) + \lambda_2 f(x_2) + \cdots + \lambda_n f(x_n) \tag{3.2}$$

for $\lambda_i \geq 0 \; i = 1, \ldots, n \; \sum_{i=1}^{n} \lambda_i = 1$.

Convex Function

A function $f(\mathrm{x})$ is convex iff $-f(\mathrm{x})$ is concave.

Theorem 1

A differentiable function $f(\mathrm{x}) : \mathbb{R}^n \to \mathbb{R}$ is concave iff

$$f(\mathrm{x} + \Delta\mathrm{x}) \leq f(\mathrm{x}) + \Delta\mathrm{x} \cdot f_{\mathrm{x}} \tag{3.3}$$

for all $\mathrm{x}, \Delta\mathrm{x} \in \mathbb{R}^n$ and where $f_{\mathrm{x}} = (\partial f/\partial x_1, \partial f/\partial x_2, \ldots, \partial f/\partial x_n)$. It is strictly concave when the strict inequality holds.

Proof Take x and $\mathrm{x} + \Delta\mathrm{x}$ for given α.

$$\begin{aligned} f\left(\alpha\left(\mathrm{x}+\Delta \mathrm{x}\right)+(1-\alpha)\mathrm{x}\right) &\geq \alpha f\left(\mathrm{x}+\Delta\mathrm{x}\right)+(1-\alpha)f(\mathrm{x}) \\ f(\mathrm{x}+\alpha\Delta\mathrm{x})-f(\mathrm{x}) &\geq \alpha(f(\mathrm{x}+\Delta\mathrm{x})-f(\mathrm{x})) \\ f(\mathrm{x})+\tfrac{f(\mathrm{x}+\alpha\Delta\mathrm{x})-f(\mathrm{x})}{\alpha} &\geq f(\mathrm{x}+\Delta\mathrm{x}) \end{aligned} \tag{3.4}$$

Take $\alpha \to 0$, then we have the desired inequality.

Negative Semidefinite A matrix H is negative semidefinite if

$$\mathrm{z}'\mathrm{Hz} \leq 0 \quad \mathrm{z} \in \mathbb{R}^n \tag{3.5}$$

It is called negative definite if $\mathrm{z}'\mathrm{Hz} < 0$ for all $\mathrm{z} \neq 0 \in \mathbb{R}^n$.

Theorem 2 A symmetric matrix H is negative semidefinite iff the principal minors alternate in sign.

Theorem 3 A twice continuously differentiable function $f : A \to \mathbb{R}$ is concave (strictly concave) iff f_{xx} is negative semidefinite (negative definite).

Exercise 1 Prove theorem 3.

Three important properties of concave functions:

1. their critical points are automatically global maxima
2. the weighted sum of concave functions is a concave function
3. the level sets of a concave function have just the right shapes for consumption and production theory.

Quasi-Concavity

Quasiconcave Function A function f is *quasi-concave* if for all y the set $\{\mathrm{x}|f(\mathrm{x}) \geq f(\mathrm{y})\}$ is convex.

Here's another definition using the level set:

A function f defined on a convex set Y is quasiconcave if every upper level set of f, $Q_a = \{\mathrm{x} \in Y | f(\mathrm{x}) \geq a\}$, is convex.[1]

[1] Recall that a set A is convex if for all $a, b \in A$ and for all $\lambda \in [0, 1]$, $\lambda a + (1 - \lambda)b \in A$.

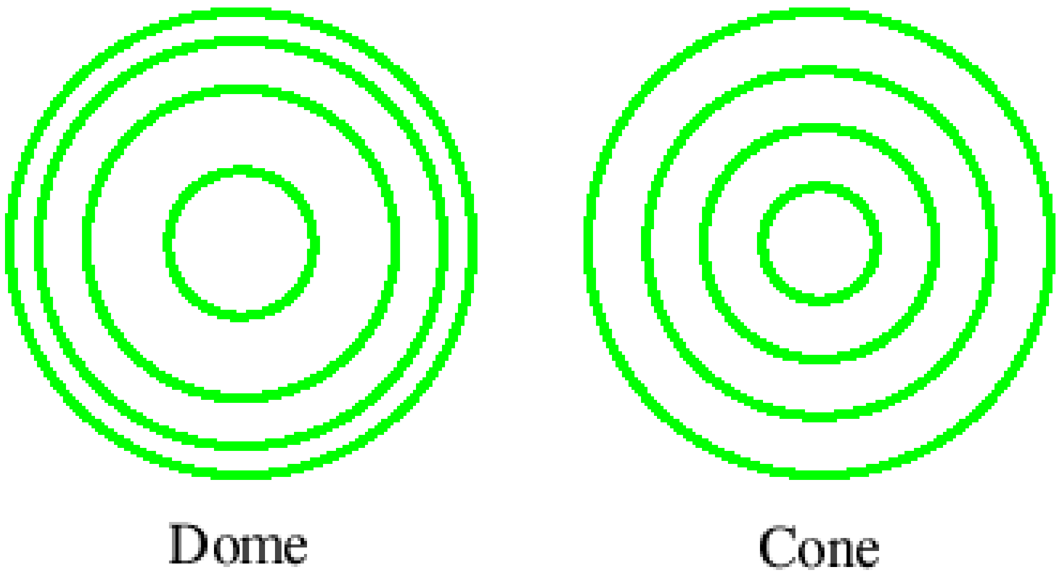

Figure 3.1: Level Sets

To determine whether a twice-differentiable function is quasiconcave or quasiconvex, we can examine the determinants of the *bordered* Hessians of the function, defined as follows:

$$D_r(\mathrm{x}) = \begin{vmatrix} 0 & f_1'(\mathrm{x}) & f_2'(\mathrm{x}) & \dots & f_r'(\mathrm{x}) \\ f_1'(\mathrm{x}) & f_{11}''(\mathrm{x}) & f_{12}''(\mathrm{x}) & \dots & f_{1r}''(\mathrm{x}) \\ f_2'(\mathrm{x}) & f_{21}''(\mathrm{x}) & f_{22}''(\mathrm{x}) & \dots & f_{2r}''(\mathrm{x}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_r'(\mathrm{x}) & f_{r1}''(\mathrm{x}) & f_{r2}''(\mathrm{x}) & \dots & f_{rr}''(\mathrm{x}) \end{vmatrix} \tag{3.6}$$

Notice that a function of n variables has n bordered Hessians $D_1, \dots, D_n$.

Let f be a function of n variables with continuous partial derivatives of first and second order in an open convex set S.

f is quasiconcave iff $D_1(\mathrm{x}) \leq 0, D_2(\mathrm{x}) \geq 0, \dots, D_n(\mathrm{x}) \leq 0$ if n is odd and $D_n(\mathrm{x}) \geq 0$ if n is even, for all x in S.

f is quasiconvex iff $D_k(\mathrm{x}) \leq 0$ for all k, for all x in S.

Strict Quasiconcavity

A function $f : D \to \mathbb{R}$ is *strictly quasiconcave* iff, for all $\mathrm{x} \neq \mathrm{y}$ in D, $f(\mathrm{x}^t) > \min[f(\mathrm{x}), f(\mathrm{y})]$ for all $t \in (0, 1)$.

Exercise 2

Let $y = x_1^\alpha x_2^\beta$. The bordered Hessian determinant is:

$$|H_\beta| = \begin{vmatrix} \alpha(\alpha-1)x_1^{\alpha-2} & x_2^\beta & \alpha\beta x_1^{\alpha-1}x_2^{\beta-1} & \alpha x_1^{\alpha-1}x_2^\beta \\ \alpha\beta x_1^{\alpha-1} & x_2^{\beta-1} & \beta(\beta-1)x_1^\alpha x_x^{\beta-2} & \beta x_1^\alpha x_2^{\beta-1} \\ \alpha x_1^{\alpha-1} & x_2^\beta & \beta x_1^\alpha x_2^{\beta-1} & 0 \end{vmatrix} \geq 0 \tag{3.7}$$

Now marginal products nonnegative require $\alpha, \beta \geq 0$. Calculate $|H_\beta|$ and show that this is sufficient for quasiconcavity of the production function.

Hint

For Cobb-Douglas, it helps to get rid of as many posers of x, notationally, as you can. for example:

$$\alpha x_1^{\alpha-1}x_2^\beta = \alpha y/x_1$$
$$\alpha(\alpha-1)x_1^{\alpha-2}x_2^\beta = \alpha(\alpha-1)y/x_1^2$$

Every concave function is quasiconcave.

Any monotonic transformation of a concave function is quasiconcave.

NB

Quasiconcave functions lack the first two advantages of concave functions. For example, consider $y = x^3$. And what about $f_1(x) = x^3$ and $f_2(x) = -x$ and $f_3 = x^3 - x$?

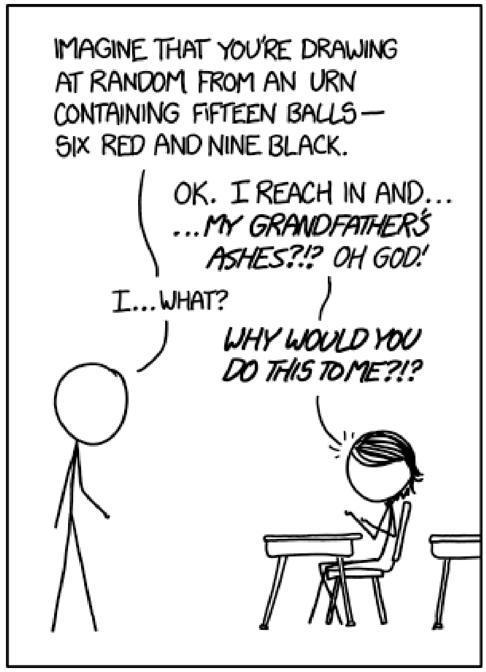

Figure 3.2: Urn (xkcd.com)

4 Production Functions

**Jehle and Reny 3*
Rubenstein Lecture 7
MWG 5.B

Simplifying Assumptions

To simplify things, we are going to make three assumptions about production functions.

The production function, $f : \mathbb{R}^n_+ \to \mathbb{R}_+$, is *continuous*, *strictly increasing*, and *strictly quasiconcave* on $\mathbb{R}^n_+$, and $f(\mathbf{0}) = 0$.

Continuity

Small changes in the vector of inputs lead to small changes in the amout of output produced.

A function $f : \mathbb{R} \to \mathbb{R}$ is continuous at a point x^0 if, for all $\varepsilon > 0$ there exists a $\delta > 0$ such that $d(x, x^0) < \delta$ implies that $d\left(f(x), f(x^0)\right) < \varepsilon$. A function is called a *continuous function* if it is continuous at every point in its domain.

Strictly Increasing

Employing strictly more of every input results in strictly more output.

Let $f : D \to \mathbb{R}$, where D is a subset of $\mathbb{R}^n$. Then f is *increasing* if $f(\mathrm{x}^0) \geq f(\mathrm{x}^1)$ whenever $\mathrm{x}^0 \geq \mathrm{x}^1$. If, in addition, the inequality is strict whenever $\mathrm{x}^0 \gg \mathrm{x}^1$, then we say that f is *strictly increasing*. If, instead, $f(\mathrm{x}^0) > f(\mathrm{x}^1)$ whenever x^0 and x^1 are distinct and $\mathrm{x}^0 \geq \mathrm{x}^1$, then we say that f is *strongly increasing*.

Strict Quasiconcavity

Mostly for convenience; implies some *complementarities* in production. See "convexity" in the notes on production sets for more details & implications.

A function $f : D \to \mathbb{R}$ is *strictly quasiconcave* iff, for all $\mathrm{x}^1 \neq \mathrm{x}^2$ in D, $f(\mathrm{tx}^1 + (1-\mathrm{t})\mathrm{x}^2) > \min\left[f(\mathrm{x}^1), f(\mathrm{x}^2)\right]$ for all $t \in (0, 1)$. Note that this is a general definition that does not rely on the differentiability of $F(\mathrm{x})$.

Theorem 1 Suppose that $f(\cdot)$ is the production function associated with a single-output technology, and let Y be the production set of this technology. Y is convex iff $f(\mathrm{x})$ is concave.

Exercise 1 Prove theorem 1.

Definitions and Properties

Marginal Product

The *marginal product* of input i is the partial derivative $\partial f(\mathrm{x})/\partial x_i$, of a (differentiable) production function. It gives the rate at which output changes per additional unit of input i employed.

Isoquant

The set of input vectors producing y units of output is called the y-level isoquant. It is a level set of f, and it is denoted by

$$Q(y) \equiv \{\mathrm{x} \geq 0 | f(\mathrm{x}) = y\}. \tag{4.1}$$

MRTS

The *marginal rate of technical substitution (MRTS)* measures the rate at which one input can be substituted for another without changing the amount of output produced. The MRTS of input j for input i is

$$MRTS_{ij}(\mathrm{x}) = \frac{\partial f(\mathrm{x})/\partial x_i}{\partial f(\mathrm{x})/\partial x_j}. \tag{4.2}$$

It is the absolute value of the slope of the isoquant through x^1 at the point x^1, an arbitrary point on the isoquant.

If $f(\mathrm{x})$ is differentiable, when $\mathrm{y} = f(x_1, x_2)$ we could use the familiar declining marginal rate of technical substitution formula

$$\frac{\partial f(\mathrm{x})/\partial x_2}{\partial f(\mathrm{x})/\partial x_1}\Big|_{yfixed} \geq 0 \tag{4.3}$$

for quasiconcavity of $f(\mathrm{x})$ or convexity of $v(\mathrm{y}) = (\mathrm{x}|\mathrm{y} \leq f(\mathrm{x}))$.

Exercise 2 Take a two-input production function $y = f(x_1, x_2)$. Derive the expression for the MRTS. Then derive the expression for the change in the MRTS with respect to x_1. Demonstrate that this is equivalent to the determinantal, i.e., the determinant of the bordered Hessian, condition for quasiconcavity.

Homogeneous Functions

Homogeneous functions feature prominantly in production economics. They simplify things and seem to be emprically valid in many cases.

Definition A function $y = f(x_1, x_2, \ldots, x_n)$ is positively homogeneous of degree λ iff

$$f(tx_1, tx_2, \ldots, tx_n) = t^\lambda f(x_1, x_2, \ldots, x_n) \qquad t > 0. \tag{4.4}$$

Properties of Homogeneous Functions

Property 1 A function homogeneous of degree zero, $f(t\mathrm{x}) = t^0 f(\mathrm{x}) = f(\mathrm{x})$, can be written by normalizing on a variable.

Example Let $t = 1/x_1$, then

$$\begin{aligned} f(x_1/x_1, x_2/x_1, \ldots, x_n/x_1) &= \left(\tfrac{1}{x_1}\right)^0 f(\mathrm{x}) = f(\mathrm{x}) \\ f(1, x_2/x_1, \ldots, x_n/x_1) &= \left(\tfrac{1}{x_1}\right)^0 f(\mathrm{x}) = f(\mathrm{x}) \\ \tilde{f}(x_2/x_1, \ldots, x_n/x_1) &= \left(\tfrac{1}{x_1}\right)^0 f(\mathrm{x}) = f(\mathrm{x}) \end{aligned} \tag{4.5}$$

where $\tilde{f}(\cdot)$ is the normalized function.

Property 2 A function homogeneous of degree one can always be written as a variable times the normalized function in Property 1.

Proof In class

Property 3 A function, homogeneous of degree λ has partial derivatives homogenous of degree $\lambda - 1$.

Proof In class

Corollary A production function homogeneous of degree 1 has marginal products which only depend on input ratios.

Property 4 A function is homogeneous of degree λ iff

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i} x_i = \lambda f(\mathrm{x}) \tag{4.6}$$

Homogeneous Production Functions …

…Are Concave

Let $f(\mathrm{x})$ be a production function satisfying the production function assumptions above and suppose it is homogeneous of degree $\alpha \in (0,1]$. Then $f(\mathrm{x})$ is a concave function of x.

Exercise 3 Prove the above statement.

What follows is the answer to exercise 3, for lecturer's and grader's notes

Answer From Jehle and Reny 2001 p 132.

Assume the production function is homogeneous of degree 1, i.e., $\alpha = 1$. Take any $\mathrm{x}_1 >> 0$ and $\mathrm{x}_2 >> 0$. Let $y_1 - f(\mathrm{x}_1$ and $y_2 - f(\mathrm{x}_2$. Since $f(\mathbf{0}) = 0$ and $f(\cdot)$ is strictly increasing, $y_1, y_2 > 0$. Since $f(\cdot)$ is homogeneous of degree 1, $f\left(\frac{\mathrm{x}_1}{y_1}\right) = f\left(\frac{\mathrm{x}_1}{y_1}\right) = 1$. Because f is quasiconcave

$$f\left(\frac{t\mathrm{x}_1}{y_1} + (1-t)\frac{\mathrm{x}_2}{y_2}\right) \geq 1 \tag{4.7}$$

for all $t \in [0,1]$.

Now choose $t^* = y_1/(y_1+y_2)$ and $(1-t^*) = y_2/(y_1+y_2)$. From eq (4.7)

$$f\left(\frac{\mathrm{x}_1}{y_1+y_2} + \frac{\mathrm{x}_2}{y_1+y_2}\right) \geq 1. \tag{4.8}$$

By linear homogeneity of f we can write

$$\frac{1}{y_1+y_2} f\left(\mathrm{x}_1 + \mathrm{x}_2\right) \geq 1. \tag{4.9}$$

and

$$f(\mathrm{x}_1 + \mathrm{x}_2) \geq y_1 + y_2 = f(\mathrm{x}_1) + f(\mathrm{x}_2). \tag{4.10}$$

This shows that eq (4.10) holds for $\mathrm{x}_1 >> 0$ and $\mathrm{x}_2 >> 0$, and because f is continuous it also holds for all $\mathrm{x}_1, \mathrm{x}_2 \geq 0$.

Now take any arbitrary $t \in [0, 1]$. Since we have assumed $\alpha = 1$, we can use eq (4.10) to write

$$f(t\mathrm{x}_1 + (1-t)\mathrm{x}_2) \geq tf(\mathrm{x}_1) + (1-t)f(\mathrm{x}_2), \tag{4.11}$$

which shows that a linear homogenous function is concave.

Now suppose f is homogeneous of degree $\alpha \in (0, 1]$. Then $f^{\frac{1}{\alpha}}$ is homogeneous of degree 1 and concave, by what we've just proven. Since $\alpha \leq 1$, $f = \left[f^{\frac{1}{\alpha}}\right]^{\alpha}$ is concave.

…Have MRTS That Only Depend On Input Ratios

For a production function that is homogeneous of degree λ, the marginal rate of technical substitution is independent of input levels and only depends on input ratios.

Separable Production Functions

In general, the MRTS between any two inputs depends on the amounts of all inputs employed. Empirically, however, researchers classify inputs into a relatively small number of types, with degrees of substitutability between those of a given type differing systematically from the degree of substitutability between those of different types. Production functions of this variety are called separable, and there are at least two major forms of separability.

Definition Let $N = \{1, \ldots, n\}$ index the set of all inputs, and suppose that these inputs can be partitioned into $S > 1$ mutually exclusive and exhaustive subsets, $N_1, \ldots, N_S$. The production function is called weakly separable if the MRTS between two inputs within the same group is independent of inputs used in other groups:

$$\frac{\partial\left(f_i(\mathrm{x})/f_j(\mathrm{x})\right)}{\partial x_k} = 0 \qquad \forall \quad i, j \in N_S \text{ and } k \notin N_S, \tag{4.12}$$

where f_i and f_j are the marginal products of inputs i and j. When $S > 2$, the production function is called strongly separable if the MRTS between two inputs from any two groups, including from the same group, is independent of all inputs outside those two groups:

$$\frac{\partial \left(f_i(\mathrm{x})/f_j(\mathrm{x})\right)}{\partial x_k} = 0 \qquad \forall \quad i \in N_S, j \in N_t \text{ and } k \notin N_S \cup N_t. \tag{4.13}$$

The Elasticity of Substitution

Definition For a production function $f(\mathrm{x})$, the elasticity of substitution of input j for input i at the point $\mathrm{x}^0 \in \mathbb{R}^n_{++}$ is defined as

$$\sigma_{ij}(\mathrm{x}^0) \equiv \left(\frac{d \ln MRTS_{ij}(\mathrm{x}(r))}{d \ln r} |_{r = x_j^0/x_i^0} \right)^{-1}, \tag{4.14}$$

where $\mathrm{x}(r)$ is the unique vector of inputs $\mathrm{x} = (x_1, \ldots, x_n)$ such that

1. $x_j/x_i = r$
2. $x_k = x_k^0$ for $k \neq i, j$
3. $f(\mathrm{x}) = f(\mathrm{x}^0)$.

The elasticity of substitution $\sigma_{ij}(\mathrm{x}^0)$ is a measure of the curvature of the $i-j$ isoquant through x^0 at x^0. When the production function is quasiconcave, the elasticity of substitution can never be negative, so $\sigma_{ij} \geq 0$. In general, the closer it is to zero, the more 'difficult' is substitution between the inputs; the larger it is, the 'easier' is substitution between them.

Returns to Scale and Varying Proportions

In the *short run*, the period of time in which at least one input is fixed, output can be varied only by changing the amounts of some inputs but not others. As amounts of the variable inputs are changed, the proportions in which fixed and variable inputs are used are also changed. *Returns to variable proportions* refer to how output responds in this situation.

Elementary measures of returns to varying proportions include the marginal product and the average product of each input. The output elasticity of input i, measuring the percentage response of output to a 1 per cent change in input i, is given by

$$\mu_i(\mathrm{x}) \equiv \frac{f_i(\mathrm{x}) x_i}{f(\mathrm{x})} = \frac{MP_i(\mathrm{x})}{AP_i(\mathrm{x})} \tag{4.15}$$

Each of these is a local measure, defined at a point.

In the long run, the firm is free to vary all inputs, and classifying production functions by their *returns to scale* is one way of describing how output responds in this situation. Specifically, returns to scale refer to how output responds when all inputs are varied in the same proportion, i.e., when the entire *scale* of operation is increased or decreased proportionally.

Returns to Scale

A production function $f(\mathrm{x})$ has the property of:

1. Constant returns to scale if $f(t\mathrm{x}) = tf(\mathrm{x})$ for all $t > 0$ and all x
2. Increasing returns to scale if $f(t\mathrm{x}) > tf(\mathrm{x})$ for all $t > 1$ and all x
3. Decreasing returns to scale if $f(t\mathrm{x}) < tf(\mathrm{x})$ for all $t > 1$ and all x

Note Notice from these global definitions of returns to scale that a production function has constant returns if it is a (positive) linear homogeneous function. Notice carefully, however, that every homogeneous production function of degree greater (less) than one must have increasing (decreasing) returns, though the converse need not hold.

Scale Elasticity For a production function $y = f(\mathrm{x})$ the scale elasticity for positive x is

$$\epsilon(\mathrm{x}) = \sum_{i=1}^{n} \frac{\partial \ln f(\mathrm{x})}{\partial \ln x_i} \tag{4.16}$$

Thus, for homogeneous production functions,

$$\epsilon(\mathrm{x}) = \lambda \tag{4.17}$$

independent of x.

Exercises 4–6 For each of the following statements, indicate whether it is true, false or uncertain and explain your answer ("uncertain" means there isn't enough information to determine the truth of the statement; it does not mean you are uncertain about the answer).

4. A firm uses 10 units of labor and 20 units of capital to produce 10 units of output. The marginal product of labor is 0.5. If there are constant returns to scale, the marginal product of capital must be 0.25.

5. If the average product of labor is at a maximum, it must equal the marginal product.

6. In the case of a production function subject to constant returns to scale, diminishing marginal productivity implies that more of one input raises the marginal product of the other.

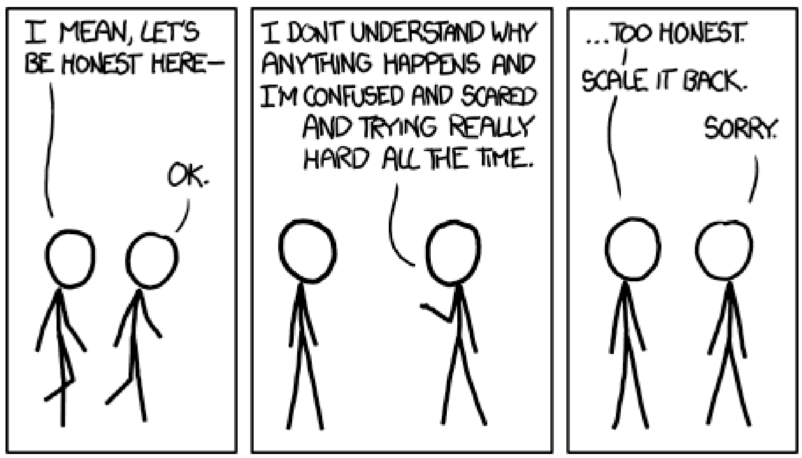

Figure 4.1: Honest (xkcd.com)

5 Costs

**Jehle and Reny 3.3*
Rubenstein Lecture 6
MWG 5.C

We start with the cost function because cost-minimization applies to all firms, whether they are in a perfectly competitive market or not.

Cost Minimization

Although we make no assumptions on the competitive environment into which the firm sells what it produces, we will assume throughout that firms are perfectly competitive on their *input* markets and that therefore they face fixed input prices.

Simple Example

Let's start with a simple example.

$$\min_{L,K} wL + rK \qquad s.t. f(K, L) = y \tag{5.1}$$

Choose L & K to produce output y given input prices w & r so that costs are minimized, i.e., lowest possible isocost curve.

$$\begin{aligned} wL + rK &= c \\ K &= \tfrac{c}{r} - \tfrac{w}{r}L \end{aligned} \tag{5.2}$$

Isocost curves:

At the minimum cost, the slope of the isoquant equals the slope of the isocost.

$$\frac{MP_L}{MP_K} = \frac{w}{r} \tag{5.3}$$

In words, this says the rate at which K is traded for L in the production process must equal the rate K & L are traded in the market.

More generally, this conclusion can be reached using Lagrange's method:

$$L = wL + rK + \lambda[y - f(L, K)] \tag{5.4}$$

Taking the first-order conditions and rearranging gives the Lagrange conditions,

$$\begin{aligned} w &= \lambda \frac{\partial f}{\partial L} \\ r &= \lambda \frac{\partial f}{\partial K} \end{aligned} \tag{5.5}$$

which can be combined to yield equation (5.3).

More Inputs

Let $\mathrm{w} = (w_1, \ldots, w_n) \geq 0$ be a vector of prevailing market prices at which the firm can buy inputs $\mathrm{x} = (x_1, \ldots, x_n)$[1]. Because the firm is a profit maximizer, which are the only preferences we are allowing firms to have—see Rubinstein(2012) it will choose to produce some level of output while using that input vector requiring the smallest money outlay.

The cost function, defined for all input prices $\mathrm{w} \gg 0$ and all output levels $y \in f(\mathbb{R}^n_+)$ is the minimum-value function,

$$c(\mathrm{w}, y) \equiv \min_{\mathrm{x} \in \mathbb{R}^n_+} \mathrm{w} \cdot \mathrm{x} \qquad s.t. \qquad f(\mathrm{x}) \geq y. \tag{5.6}$$

If $\mathrm{x}(\mathrm{w}, y)$ solves the cost-minimization problem, then

$$c(\mathrm{w}, y) = \mathrm{w} \cdot \mathrm{x}(\mathrm{w}, y). \tag{5.7}$$

Note Because f is strictly increasing, the constraint will always be binding at a solution.

Exercise Show that cost minimization implies that the marginal rate of technical substitution between any two inputs is equal to the ratio of their prices.

The solution to the cost minimization problem is a function of w and y. Assume $\mathrm{w} \gg 0$ and f is strictly quasiconcave. Then $\mathrm{x}^* \equiv \mathrm{x}(\mathrm{w}, y)$ is the vector of inputs that minimize the cost of producing y units of output at input prices w. This is the firm's *conditional input demand* because it is conditional on the level of output y.

In-class Exercise Suppose the firm's technology is the two-input CES form. Find the conditional input demands and the cost function.

[1] I will be using the notation of Jehle and Reny (2001), which differs from that of Mas-Colell, Whinston, and Green (1995). Instead of x, MWG use z.

Properties of the Cost Function

Here we follow the nomenclature of Jehle and Reny (2001), which is very similar to Mas-Colell, Whinston, and Green (1995), but not identical.

Cost Fn Properties

If f is continuous and strictly increasing then $c(\mathrm{w}, y)$ is

1. Zero when $y = 0$,
2. Continuous on its domain,
3. For all $\mathrm{w} \gg 0$, strictly increasing and unbounded above in y,
4. Increasing in w,
5. Homogeneous of degree one in w,
6. Concave in w,
7. And if f is strictly quasiconcave, $c(\mathrm{w}, y)$ is differentiable in w at (w^0, y^0) whenever $\mathrm{w}^0 \gg 0$, and

$$\frac{\partial c(\mathrm{w}^0, y^0)}{\partial w_i} = x_i(\mathrm{w}^0, y^0), \qquad i = 1, \ldots, n. \tag{5.8}$$

Proof of 3

Suppose that $c(\mathrm{w}, y)$ were not strictly increasing in y, and let x' and x'' denote optimal input vectors for required production levels y' and y'', respectively, where $y'' > y'$ and $\mathrm{w} \cdot \mathrm{x}' \geq \mathrm{w} \cdot \mathrm{x}'' > 0$. Consider an input vector $\tilde{\mathrm{x}} = \alpha \mathrm{x}''$, where $\alpha \in (0, 1)$. By continuity of $f(\cdot)$, there exists an α close enough to 1 such that $f(\tilde{\mathrm{x}}) > y'$ and $\mathrm{w} \cdot \mathrm{x}' > \mathrm{w} \cdot \tilde{\mathrm{x}}$. But this contradicts x' being optimal in the CMP with required production level y'.

Proof of 4

Suppose that the input price vectors w'' and w' have $\mathrm{w}''_l \geq \mathrm{w}'_l$ and $\mathrm{w}''_k = \mathrm{w}'_k$ for all $k \neq l$. Let x'' be an optimizing vector in the CMP for input prices w''. Then $\mathrm{c}(\mathrm{w}'', y) = \mathrm{w}'' \cdot \mathrm{x}'' \geq \mathrm{w}' \cdot \mathrm{x}'' \geq c(\mathrm{w}', y)$, where the latter inequality follows from the definition of $c(\mathrm{w}', y)$.

Proof of 5

The constraint set of the CMP is unchanged when input prices change. Thus, for any scalar $\alpha > 0$, minimizing $(\alpha \mathrm{w}) \cdot \mathrm{x}$ on this set leads to the same optimal production as minimizing $\mathrm{w} \cdot \mathrm{x}$. Letting x^* be optimal in both circumstances, we have $c(\alpha \mathrm{w}, y) = \alpha \mathrm{w} \cdot \mathrm{x}^* = \alpha c(\mathrm{w}, y)$.

Proof of 6 For concavity, fix a required output level $\bar{y}$, and let $\mathrm{w}'' = \alpha\mathrm{w}+(1-\alpha)\mathrm{w}'$ for $\alpha \in [0,1]$. Suppose that x'' is an optimal input vector in the CMP when input prices are w''. If so,

$$\begin{aligned} c(\mathrm{w}'', \bar{y}) &= \mathrm{w}'' \cdot \mathrm{x}'' \\ &= \alpha \mathrm{w} \cdot \mathrm{x}'' + (1-\alpha)\mathrm{w}' \cdot \mathrm{x}'' \\ &\geq \alpha c(\mathrm{w}, \bar{y}) + (1-\alpha(\mathrm{w}', \bar{y}). \end{aligned} \tag{5.9}$$

where the last inequality follows because $f(\mathrm{x}'') \geq \bar{y}$ and the definition of the cost function imply that $\mathrm{w} \cdot \mathrm{x}'' \geq c(\mathrm{w}, \bar{y})$ and $\mathrm{w}' \cdot \mathrm{x}'' \geq c(\mathrm{w}', \bar{y})$. when $x\mathrm{d} >> 0$

Homothetic Tidbits

Definition *Homothetic*

A real-valued function h on $D \subset \mathbb{R}^n$ is called homothetic if it can be written in the form $g\,(f(\mathrm{x}))$, where $g : \mathbb{R} \to \mathbb{R}$ is strictly increasing and $f : D \to \mathbb{R}$ is homogeneous of degree 1.

Theorem A production function $f(\cdot)$ is *homothetic* iff the MRTS is only dependent on input ratios rather than scale.

Definition Technology is homothetic if $y \leq F\,(f(\mathrm{x}))$, $f\,(t\mathrm{x}) = tf(\mathrm{x}) \quad t > 0, F' > 0$.

Theorem The cost function is of the form

$$c\,(\mathrm{w}, y) = h(y)\phi(\mathrm{w}) \tag{5.10}$$

ϕ linearly homogeneous, continuous on $\mathbb{R}^n_{++}$, and concave if and only if technology is homothetic.

Proof In class.

Conditional Input Demand Properties

If the production function satisfies the Assumptions and the associated cost function is twice continuously differentiable, then

1. $\mathrm{x}(\mathrm{w}, y)$ is homogeneous of degree zero in w,

2. The substitution matrix, defined and denoted

$$\sigma^*(\mathrm{w}, y) = \begin{bmatrix} \frac{\partial x_1(\mathrm{w},y)}{\partial w_1} & \cdots & \frac{\partial x_1(\mathrm{w},y)}{\partial w_n} \\ \vdots & & \vdots \\ \frac{\partial x_n(\mathrm{w},y)}{\partial w_1} & \cdots & \frac{\partial x_n(\mathrm{w},y)}{\partial w_n} \end{bmatrix}, \tag{5.11}$$

is symmetric and negative semi-definite. In particular, the negative semi-definiteness property implies that $\partial x_i(\mathrm{w}, y)/\partial w_i \leq 0$ for all i.

Cost and Conditional Input Demands when Production is Homothetic

Theorem

1. When the production function satisfies the Assumptions and is homothetic, the conditional input demands are multiplicitivly separable in input prices and output and can be written $\mathrm{x}(\mathrm{w}, y) = h(y)\mathrm{x}(\mathrm{w}, 1)$, where $h'(y) > 0$ and $\mathrm{x}(\mathrm{w}, 1)$ is the conditional input demand for 1 unit of output.

2. When the production function is homogeneous of degree $\alpha > 0$,

   a) $c(\mathrm{w}, y) = y^{1/\alpha} c(\mathrm{w}, 1)$

   b) $\mathrm{x}(\mathrm{w}, y) = y^{1/\alpha} \mathrm{x}(\mathrm{w}, 1)$.

Exercise Prove the previous two propositions.

Implications

1. Conditional demand function ratios are independent of output.

2. Cost shares are independent of output.

3. $F^{-1}(y)$ for homogeneous functions is $y^{1/k}$ where k is the degree of homogeneity.

4. What about the nature of elasticities of substitution?

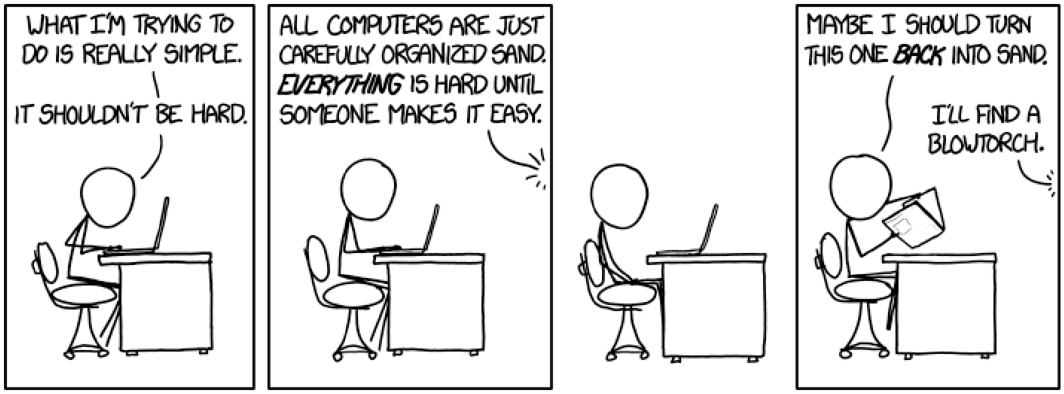

Figure 5.1: Shouldn't be hard (xkcd.com)

6 Profit Maximization

**Jehle and Reny 3.5*
Rubenstein Lecture 6
MWG 5.C


THEY'RE THE ONES GIVING CHIPS AWAY!
IF THEY DON'T SEE THE ARBITRAGE POTENTIAL, SUCKS FOR THEM.
CHIPS
IN A DEEP SENSE, SOCIETY FUNCTIONS ONLY BECAUSE WE GENERALLY AVOID TAKING THESE PEOPLE OUT TO DINNER.

Part III

Consumer Theory—A Modern Approach

7 Preferences

**Rubenstein Lecture 1*
MWG 1
Kreps 2
Jehle and Reny 1

Useful Tools

Definitions

Relation Properties

Suppose R is a relation on a set A.

1. Relation R is *reflexive* if $\mathrm{x}R\mathrm{x}$ for every $\mathrm{x} \in A$
2. Relation R is *symmetric* if $\mathrm{x}R\mathrm{y}$ implies $\mathrm{y}R\mathrm{x}$ for all $\mathrm{x}, \mathrm{y} \in A$.
3. Relation R is *transitive* if whenever $\mathrm{x}R\mathrm{y}$ and $\mathrm{y}R\mathrm{z}$, then also $\mathrm{x}R\mathrm{z}$.

Domain, Codomain, Range

For a function $f : A \to B$, the set A is called the *domain* of f. (Think of the domain as the set of possible "input values" for f.) The set B is called the *codomain* of f. The *range* of f is the set $\{f(a) : a \in A\} = \{b : (a, b) \in f\}$. (Think of the range as the set of all possible "output values" for f. Think of the codomain as a sort of "target" for the outputs.)

NB

The range is a subset of the codomain, but it does not necessarily equal the codomain.

Injective, Surjective

A function $f : A \to B$ is:

1. injective (or one-to-one) if for every $\mathrm{x}, \mathrm{y} \in A$, $\mathrm{x} \neq \mathrm{y}$ implies $f(\mathrm{x}) \neq f(\mathrm{y})$
2. surjective (or onto) if for every $\mathrm{b} \in B$ there is an $\mathrm{a} \in A$ with $f(\mathrm{a}) = \mathrm{b}$; (A function is surjective if and only if its codomain equals its range.)
3. bijective if f is both injective and surjective.

Ultimate Objective

Our goal is to develop a formal theory of choice.

Note Ask yourself why it is important to develop a formal theory of choice. In other words, why do we care? If you can't come up with a good answer that really *satisfies* you—and I mean *really satisfies* you—then I recommend that you quit this program, backpack through Europe, climb the Inca trail to Machu Pichu, or make the journey to Llasa, and figure out what does *really satisfy* you.

If it is extremely clear to you why developing a formal theory of choice is important, then we will begin by looking at *preferences*, independent of choice (we'll get to choice soon enough).

Behavioral Restrictions

Measurement

Preferences Preferences are the mental attitude of an individual toward alternatives.

Measuring preferences is not easy. In fact, if we are going to build a formal theory about preferences, we need to make a lot of *heroic* assumptions. To illustrate, we will use two questionnaires to define a *preference relation*. The first questionnaire, which we will call the *Q questionnaire*, more clearly illustrates the assumptions we are making and the restrictions we are imposing. These "restrictions" will make it possible to develop a formal theory of preferences. These restrictions also make it *impossible* to develop a formal theory of *all* behavior, as we will see. The formal theory we develop will only apply to behavior that fits into the special restrictions we impose here, at the very beginning.

The simplest way to define a preference relation is to look at only two goods (or bundles of goods) at a time. So our entire theory of consumer behavior will be based on the following questionnaires, which fundamentally compare just two (bundles of) goods at a time, x and y.

Questionnaire Q

$Q(\text{x}, \text{y})$ (for all distinct x and y in X (more on X later)):

How do you compare x and y? Choose one and only one of the following three options:

- ☐ I prefer x to y.
- ☐ I prefer y to x.

☐ I am indifferent.

Everyone has to answer this questionnaire for any and every two (bundles of) goods by choosing one, and only one, answer. They cannot choose more than one answer and cannot abstain.

$$X = \left\{ \begin{array}{ll} \text{get up} & \text{sleep in} \\ \text{jeans} & \text{chinos} \\ \text{cinnamon roll} & \text{cereal} \end{array} \right\} \xrightarrow{Q(\mathrm{x,y})} \left\{ \begin{array}{l} \text{I prefer x to y.} \\ \text{I prefer y to x.} \\ \text{I am indifferent.} \end{array} \right\} \xrightarrow{f(Q(\mathrm{x,y}))} \left\{ \begin{array}{c} \succ \\ \prec \\ I \end{array} \right\}$$

Implied restrictions

The above questionnaire, along with the rule to choose exactly one answer, might seem innocuous. In fact, it places a lot of restrictions on the type of behavior we can study. For example, it rules out:

1. a lack of ability to compare, such as
   - ☐ They are incomparable.
   - ☐ I don't know what x is.
   - ☐ I have no opinion.
   - ☐ I both prefer x over y and y over x.
2. a dependence on other factors, such as
   - ☐ It depends on what my parents think.
   - ☐ It depends on the circumstances (sometimes I prefer x, but usually I prefer y).
3. and most importantly, intensity of preferences, such as
   - ☐ I somewhat prefer x.
   - ☐ I love x and I hate y.

Assumptions

A person is capable of comparing *any* two things, and we are ignoring the intensity of preferences.

Defining preferences

We've started by talking about x and y without really defining them. First, think about making a decision–any decision–and group together all of the possible choices you have related to this decision. We're going to call groups of things, anything, a *set*. And we're going to call the group of choices related to your decision your *choice set*. We will represent your choice set with the letter X, capital 'x'. Now we can define x and y as elements of X.

We are going to represent answers to the question $Q(\mathrm{x}, \mathrm{y})$ with the function f, which assigns exactly one of three "values" to any pair (x, y) of distinct elements in X. These three potential values are:

1. $\mathrm{x} \succ \mathrm{y}$ (x is preferred to y)
2. $\mathrm{y} \succ \mathrm{x}$ (y is preferred to x)
3. I, which stands for indifference.

$f(\mathrm{x}, \mathrm{y})$ is the answer to the question $Q(\mathrm{x}, \mathrm{y})$.

More Restrictions

Not all answers to $Q(\mathrm{x}, \mathrm{y})$ count as *preferences over the set X*. We're going to make some more restrictions. The restrictions we are about to make form the foundation of what we will refer to as *rational* behavior. In other words, behavior violating the following restrictions will be deemed *irrational*.

Definition 1 *Preferences on a set X* are a function f that assigns to any pair (x, y) of distinct elements in X exactly one of the three "values" $\mathrm{x} \succ \mathrm{y}$, $\mathrm{y} \succ \mathrm{x}$, or I so that for any three different elements x, y and z in X, the following two properties hold:

- *No order effect*: $f(\mathrm{x}, \mathrm{y}) = f(\mathrm{y}, \mathrm{x})$.
- *Transitivity*:
  - if $f(\mathrm{x}, \mathrm{y}) = \mathrm{x} \succ \mathrm{y}$ and $f(\mathrm{y}, \mathrm{z}) = \mathrm{y} \succ \mathrm{z}$ then $f(\mathrm{x}, \mathrm{z}) = \mathrm{x} \succ \mathrm{z}$ and
  - if $f(\mathrm{x}, \mathrm{y}) = I$ and $f(\mathrm{y}, \mathrm{z}) = I$ then $f(\mathrm{x}, \mathrm{z}) = I$.

The Questionnaire R

The more standard theory of choice, i.e., that in Mas Colel, defines preferences using a different, but equivalent, questionnaire. Let's call this one $R(\text{x}, \text{y})$

Is x at least as preferred as y?

- ☐ Yes
- ☐ No

1. The respondent must choose one, and only one, answer.
2. The answer to at least one of the questions $R(\text{x}, \text{y})$ and $R(\text{y}, \text{x})$ must be Yes.
3. For every $\text{x}, \text{y}, \text{z} \in X$, if the answers to the questions $R(\text{x}, \text{y})$ and $R(\text{y}, \text{z})$ are Yes, then so is the answer to the question $R(\text{x}, \text{z})$.

Exercise What is the domain of $R(\text{x}, \text{y})$? What is it's codomain? Does the codomain equal the range?

- Domain of $R(\text{x}, \text{y}) = X \times X$
- Codomain of $R(\text{x}, \text{y}) =$ {Yes, No, I don't know, incomparable, no opinion, it depends, etc.}
- Range of $R(\text{x}, \text{y}) =$ {Yes, No}

Rather than represent answers to this questionnaire with a function, we will use a specific mathematical concept called a *binary relation*. A binary relation is a concept from set theory. A binary relation on X is a subset of X^2.

The binary relation looks like this: $\succsim$.

We will use $\succsim$ to define a response to the R questionnaire as x $\succsim$ y if the answer to the question $R(\text{x}, \text{y})$ is Yes.

Analogous restrictions to those used with questionnaire Q to define "rational" behavior are:

Definition 2

Preferences on a set X is an asymmetric binary relation $\succsim$ on X satisfying:

- *Completeness*: For any $\mathrm{x}, \mathrm{y} \in X$, $\mathrm{x} \succsim \mathrm{y}$ or $\mathrm{y} \succsim \mathrm{x}$.
- *Transitivity*: For any $\mathrm{x}, \mathrm{y}, \mathrm{z} \in X$, if $\mathrm{x} \succsim \mathrm{y}$ and $\mathrm{y} \succsim \mathrm{z}$, then $\mathrm{x} \succsim \mathrm{z}$.

NB These restrictions are not imposed because they accurately describe human behavior. Instead we impose them because they are the weakest restrictions needed for a mathematically consistent theory.

Questionnaire Q is equivalent to Questionnaire R

Questionnaire Q illustrated the underlying, implicit assumptions made before we even begin to define preferences. Questionnaire R is the traditional one used to define preferences, and the one that the formal theory of choice will be based on. Showing the equivalence between these shows that the traditional theory is based on a bunch of implicit behavioral restrictions.

In other words, we want to show that saying "I prefer x to y" is equivalent to saying "I find x to be at least as good as y, and I don't find y to be at least as good as x."

Proof

$T : Q \to R$ is injective

First, we need to consider answers to Q that meet Definition 1. To show equivalence, these answers must correspond to one and only one answer to R, and that answer to R must satisfy Definition 2.

We need to show that our translation of each of the three answers to Q corresponds to unique answer(s) to R.

1. Note that order doesn't matter for Q, e.g, $Q(\mathrm{x}, \mathrm{y}) = Q(\mathrm{y}, \mathrm{x}) = \mathrm{x} \succ \mathrm{y}$. But it *does* matter for R; even if $R(\mathrm{x}, \mathrm{y})$ is Yes, $R(\mathrm{y}, \mathrm{x})$ can be No. Hence, to establish a one-to-one correspondence, we will need to consider the answers $R(\mathrm{x}, \mathrm{y})$ and $R(\mathrm{y}, \mathrm{x})$ together as a pair.
2. The table below gives an example of the correspondence we are looking for.
3. The table demonstrates that each answer to Q maps to just one pair of answers to R. Call this correspondence T.

Response to $Q(\mathrm{x}, \mathrm{y})$ and $Q(\mathrm{y}, \mathrm{x})$	Response to $R(\mathrm{x}, \mathrm{y})$	Response to $R(\mathrm{y}, \mathrm{x})$
$\mathrm{x} \succ \mathrm{y}$	Yes	No
I	Yes	Yes
$\mathrm{y} \succ \mathrm{x}$	No	Yes

To show injection, show the contrapositive: Suppose there are two answers to the question Q and the table (T) maps these two answers to the same row of responses to question R, i.e., $R(\mathrm{x}, \mathrm{y})$ and $R(\mathrm{y}, \mathrm{x})$. Based on the table ($T$), these two answers to the question Q must be the same.

T implies R is a preference relation

Next, we need to show that if T is injective and f is a preference relation, then the answers to R in the table meet Definition 2 and are also a preference relation.

1. *Completeness*: To each of the three answers to $Q(\mathrm{x}, \mathrm{y})$ and $Q(\mathrm{y}, \mathrm{x})$, at least $R(\mathrm{x}, \mathrm{y})$ or $R(\mathrm{y}, \mathrm{x})$ is yes.
2. *Transitivity*: Assume $R(\mathrm{x}, \mathrm{y})$ and $R(\mathrm{y}, \mathrm{z})$ answers are 'yes'. Then $Q(\mathrm{x}, \mathrm{y}) = \mathrm{x} \succ \mathrm{y}$ or $= I$ and $Q(\mathrm{y}, \mathrm{z}) = \mathrm{y} \succ \mathrm{z}$ or $= I$. By Q-transitivity, $Q(\mathrm{x}, \mathrm{z})$ must be $\mathrm{x} \succ \mathrm{z}$ or I. So the answer to $R(\mathrm{x}, \mathrm{z})$ is "Yes".

Now we need to show that we can go the other way.

$T : Q \to R$ is surjective

Do unique responses to R imply a response to Q? Yes, we can see that from the table.

T implies Q is a preference relation

Next we need to consider answers to R that meet Definition 2 and show that these correspond to an answer to Q that satisfies Definition 1.

If T is surjective and the answers to question R are a preference relation as by Definition 2, do answers to question Q satisfy Definition 1 via the correspondence in the table?

1. *Completeness*: By the completeness of $\succsim$, for any two elements x and y, one of the entries in the right-hand column is applicable. Since these map directly to unique answers to Q, those Q-answers are complete.
2. *Transitivity*: If $f(\mathrm{x}, \mathrm{y}) = \mathrm{x} \succ \mathrm{y}$ and $f(\mathrm{y}, \mathrm{z}) = \mathrm{y} \succ \mathrm{z}$, then
   - $\mathrm{x} \succsim \mathrm{y}$ and not $\mathrm{y} \succsim \mathrm{x}$, and
   - $\mathrm{y} \succsim \mathrm{z}$ and not $\mathrm{z} \succsim \mathrm{y}$.

Therefore, by transitivity of $\succsim$,

- x $\succsim$ z, and
- not z $\succsim$ x (since if z $\succsim$ x then the transitivity of $\succsim$ would imply z $\succsim$ y).

so $f(\mathrm{x},\mathrm{z}) = \mathrm{x} \succ \mathrm{z}$.

Figure 7.1: Research ethics (xkcd.com)

8 Choice

Rubenstein Lecture 3
MWG 1.C,D
Kreps 2.1

So far we've spoken about preferences, but not behavior. Now we're going to develop a full description of behavior in all scenarios we imagine the agent might confront in a certain *context*.

Choice Problem

Choice Sets

We will think about choice problems in terms of sets of possible options.

- X is the *grand set of possibilities*.
- A *choice problem*, A, is a non-empty subset of X. It is an exhaustive listing of all the choice experiments relevant to the current decision. (You might think of it as a budget set.)
- Behavior is only defined on a "set of subsets of X," D.
- Context is the pair (X, D).

Exercise

Consider the grand set of possibilities $X = \{a, b, c\}$. Write the set of all nonempty subsets of X.

Example

Choosing between the status quo, s, and a set of alternatives Y. $X = Y \cup s$. D is the set of all subsets of X that contain s.

Implicit Assumption

Modeling a choice scenario as a set of alternatives implies assuming that the agents' choice does not depend on the way the alternatives are presented.

Modelling Choice

Choice Function

Given a context (X, D), a *choice function*, C, assigns to each set $A \in D$ a unique element of A with the interpretation that $C(A)$ is the chosen element from the set A.

Think of $C(A)$ as the actual decision of a living, breathing person when facing choice set A.

Rationalizing

Rational: A choice function is *rational* if, given the preference relation $\succsim$ on the set X, $C(A)$ is the element of $A \in D$ that is $\succsim$-optimal.

Induced Choice Function: A choice function induced by $\succsim$, $C_{\succsim}$, is a function that assigns to every non-empty set $A \in D$ the $\succsim$-best element of A. This choice function is called the *induced choice function*.[1] The choice is induced by the preference relation $\succsim$.

NB: The preference relation is fixed; it is independent of the choice set being considered.

Think of $C_{\succsim}(A)$ like this: First, we give someone Questionnaire R for every possible pair of goods in A. If we require the person to obey the rules of completeness and transitivity, then the set of all those answers defines the person's preference relation. We could then write a computer program that would use these data to pick the most-preferred item in the set A. We call that computer program the induced choice function, $C_{\succsim}(A)$.

Rationalizing Behavior

A choice function can be *rationalized* if there is a preference relation, $\succsim$ on X so that $C = C_{\succsim}$ (i.e., $C(A) = C_{\succsim}(A)$ for any A in the domain of C). In other words, consumers behave *as if* they were maximizing a preference relation.

In general, there may be more than one rationalizing preference relation $\succsim$ for a given choice structure.

Condition for Rationalizing

What follows is a condition under which a choice function can indeed be presented as if derived from some preference relation (i.e., can be rationalized).

Condition α: For any two problems $A, B \in D$, if $A \subset B$ and $C(B) \in A$ then $C(A) = C(B)$.

In other words, expanding the set of alternatives does not affect the optimal choice.

"Condition α" is sufficient for *rationalization*

[1] Google defines *induced* as: bring about or give rise to.

Proposition

Assume that C is a choice function with a domain containing at least all subsets of X of size 2 or 3. If C satisfies Condition α then there is a preference $\succsim$ on X so that $C = C_{\succsim}$.

Proof

Define $\succsim$ by $x \succsim y$ if $x = C(\{x, y\})$.

Is this a valid preference relation? Does it satisfy *completeness*[2] and *transitivity*[3]?

Completeness Remember the definition of C, "C assigns to each set $A \in D$ a unique element of A …" That means any valid choice function will choose the same element $a \in A$. Thus $C(\{x, y\}) = x$ or $C(\{x, y\}) = y$. So, by our definition of $\succsim$ above, $x \succsim y$ or $y \succsim x$.

Transitivity We want to show that if $x \succsim y$ and $y \succsim z$ then $x \succsim z$.

Using the definition of $\succsim$, we need to show that if $C(\{x, y\}) = x$ and $C(\{y, z\}) = y$ then $C(\{x, z\}) = x$.

We will use the "proof by contradiction" method.

- Assume not: $C(\{x, z\}) \neq x$ then $C(\{x, z\}) = z$.
- By Condition α and $C(\{x, z\}) = z$, $C(\{x, y, z\}) \neq x$.
- By Condition α and $C(\{x, y\}) = x$, $C(\{x, y, z\}) \neq y$.
- By Condition α and $C(\{y, z\}) = y$, $C(\{x, y, z\}) \neq z$.
- But this contradicts $C(\{x, y, z\}) \in \{x, y, z\}$.

Therefore $C(\{x, z\}) = x$ and, by definition of $\succsim$ proves that $\succsim$ is transitive.

We have established that the relation $\succsim$ on X is a *preference* relation. All that's left to show is that consumers behave as if they were maximizing this preference relation.

[2] $x \succsim y$ or $y \succsim x$
[3] if $x \succsim y$ and $y \succsim z$ then $x \succsim z$

"as if" Show $C(B) = C_{\succsim}(B)$. Again, we will use the proof by contradiction method.

- Assume $C(B) = x$ and $C_{\succsim}(B) \neq x$ for all $x \in B$.
- By the definition of $C_{\succsim}$, $\exists\, y \in B$ such that $y \succ x$.
- By definition of $\succsim$, $y \succ x$ implies $C(\{x, y\}) \neq x$[4], and since $C(\{x, y\})$ has to equal something, it equals y.
- But, by Condition α and $C(B) = x$, $C(\{x, y\})$ must equal x for all $x, y \in B$.
- This is a contradiction, so $C_{\succsim}$ must equal x and therefore is equal to $C(B)$.

NB Note that our subtle assumption that X is nonempty is not innocuous–we rule out the possibility that the consumer is *unable* to choose among the elements of A. One might be tempted to redefine "not choosing" as yet another element of the set of alternatives, A. This causes two problems:

1. As in our example above, we are limited to choice sets in A that also contain "make no choice"
2. Suppose x and y are too difficult to rank. Then we have $C(x, y, \text{"make no choice"}) =$ "make no choice". But what happens when we consider only x? $C(x, \text{"make no choice"}) =$ x, which, by revealed preference, makes our consumer irrational.

Rationalizing Irrational Choices

Descriptive or Prescriptive?

There is a tendency to think of "rational" in the normative sense. This suggests that anyone whose behavior is technically "irrational" should change their behavior to fit the rational model. Don't forget your role as social scientists is to *describe* behavior, not *prescribe* it.

[4] Think contrapositive; $P \to Q \implies \neg Q \to \neg P$

Irrationality versus Information

Consider the following: A man at a restaurant, when faced with a choice between *steak tartare*, which is made with raw beef, and *chicken*, chooses chicken. In a parallel universe, the man faces a choice among *steak tartare*, *chicken*, and *frog legs*. Given this choice set, the man chooses steak tartare. This looks like irrational behavior. But the man explains: steak tartare requires very special preparation to prevent bacteria from infesting the meat, so I typically stay away from it except when it is prepared by the best French chefs. When I saw frog legs on the menu, I knew the chef was excellent and I could trust the steak tartare.

The man isn't irrational; the choice set is wrong. Instead of *C(steak tartare, chicken)* it should have been *C(steak tartare prepared poorly, chicken)*. And instead of *C(steak tartare, chicken, frog legs)*, it should have been *C(steak tartare prepared correctly, chicken, frog legs)*.

This leads to the following condition:

Condition An alternative, a, must have the same meaning for every A which contains a.

Choice Correspondence and Revealed Preference

We might consider expanding our concept of a *choice function*, which allows for a single optimal choice from a choice set, to a *choice correspondence*, which allows for multiple optimal choices from a choice set and, hence, indifference.

Choice Correspondence

A *choice correspondence*, C, assigns to every nonempty $A \subseteq X$ a nonempty *subset* of A, that is, $\emptyset \neq C(A) \subseteq A$.

For instance, consider the extended choice set (A, f) where f is the *frame* that accompanies the choice set A, like the default option, or the order in which alternatives are presented. We could define a *choice correspondence* as $C(A) = \{x | x = c(A, f) \text{ for some } f\}$. Define the *induced choice correspondence* as $C_{\succsim}(A) = \{x \in A | x \succsim y \text{ for all } y \in A\}$ given the preference relation $\succsim$.

When $x, y \in A$ and $x \in C(A)$, we say that x is revealed to be at least as good as y. If, in addition, $y \notin C(A)$, we say that x is revealed to be strictly better than y. Condition α is now replaced by condition WARP, which requires that if x is revealed to be at least as good as y, then y is not revealed to be strictly better than x.

The Weak Axiom of Revealed Preference (WARP)

Axiom We say that C satisfies WARP if whenever $x, y \in A \cap B$, $x \in C(A)$ and $y \in C(B)$, it is also true that $x \in C(B)$.

This implies two properties:

Condition α If $a \in A \subset B$ and $a \in C(B)$ then $a \in C(A)$.

Condition β If $a, b \in A \subset B$, $a \in C(A)$ and $b \in C(B)$, then $a \in C(B)$.

Exercise Verify that conditions α and β are equivalent to WARP for any choice correspondence with a domain satisfying that if A and B are included in the domain, then so is their intersection.

Proposition

Assume that C is a choice correspondence with a domain that includes at least all subsets of size 2 or 3. Assume that C satisfies WARP. Then there is a preference $\succsim$ so that $C = C_{\succsim}$.

Proof

The proof is up to you.

Exercise Suppose that $X = \{x, y, z\}$, $D = \{\{x, y\}, \{y, z\}, \{x, z\}\}$, $C(\{x, y\}) = \{x\}$, $C(\{y, z\}) = \{y\}$, and $C(\{x, z\}) = \{z\}$. Show that this choice correspondence satisfies WARP, as claimed by MWG Example 1.D.1.

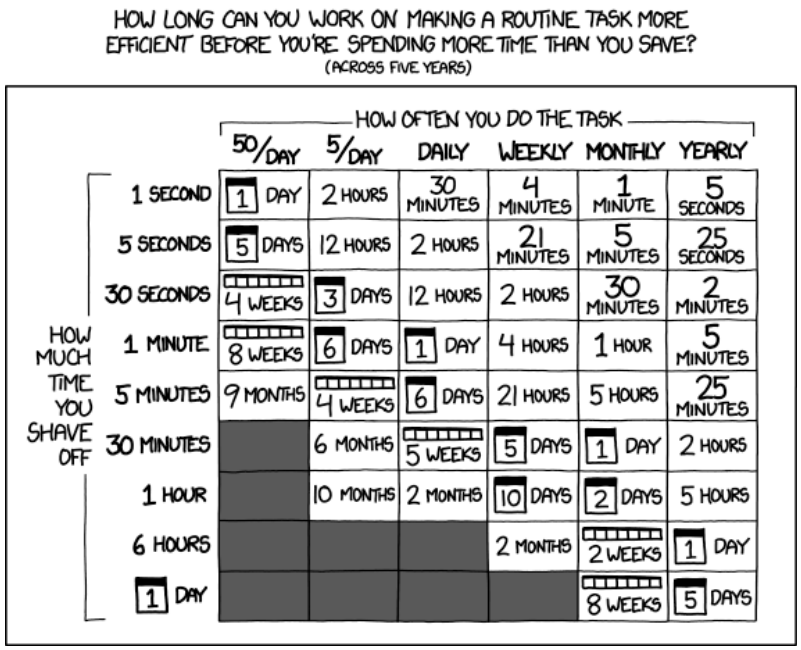

Figure 8.1: Is it worth the time? (xkcd.com)

Quiz

Framing

1. An outbreak of disease is expected to cause 600 deaths in the US. Two mutually exclusive programs are expected to yield the following results:
   a) 400 people will die.
   b) With probability 1/3, 0 people will die and with probability 2/3, 600 people will die.
2. An outbreak of disease is expected to cause 600 deaths in the US. Two mutually exclusive programs are expected to yield the following results:
   a) 200 people will be saved.
   b) With probability 1/3, all 600 will be saved and with probability 2/3, none will be saved.

78% chose 1b, but only 28% chose 2b

Similarities and Simplification

1. Choose one of the two roulette games *a* or *b*. Your prize is the one corresponding to the outcome of the chosen roulette game as specified in the following tables:

a)

Color	White	Red	Green	Yellow
Chance %	90	6	1	3
Prize $	0	45	30	-15

b)

Color	White	Red	Green	Yellow
Chance %	90	7	1	2
Prize $	0	45	-10	-15

2. A different set of roulette payoffs

a)

Color	White	Red	Green	Blue	Yellow
Chance %	90	6	1	1	2
Prize $	0	45	30	-15	-15

b)

Color	White	Red	Green	Blue	Yellow
Chance %	90	6	1	1	2
Prize $	0	45	45	-10	-15

58% chose 1, and nobody chose 3.

Reason-based Choice

1. Let $a = (a_1, a_2)$ be a holiday package of a_1 days in Paris and a_2 days in London. Choose one of the three vectors $a = (7, 4)$, $b = (4, 7)$, $c = (6, 3)$
2. Let $a = (a_1, a_2)$ be a holiday package of a_1 days in Paris and a_2 days in London. Choose one of the three vectors $a = (7, 4)$, $b = (4, 7)$, $c = (3, 6)$

The modal choice in question 1 was a, while the modal choice in question 2 was b.

Mental Accounting

1. Imagine that you have decided to see a play and paid the admission price of \$10 per ticket. As you enter the theater, you discover that you have lost the ticket. The seat was not marked, and the ticket cannot be recovered. Would you pay \$10 for another ticket?
2. Imagine that you have decided to see a play where the admission is \$10 per ticket. As you arrive at the theater, you discover that you have lost a \$10 bill. Would you still pay \$10 for a ticket for the play?

Only 46% said they would buy another ticket after losing the first one (question 1) while 88% said they would buy a ticket after losing the \$20 bill (question 2).

Discussion Questions

Problem 1

The following are descriptions of decision making procedures. Discuss whether the procedures can be described in the framework of the choice model discussed in this lecture and whether they are compatible with the “rational man” paradigm.

1. The decision maker has in mind a ranking of all alternatives and chooses the alternative that is the worst according to this ranking.
2. The decision maker chooses an alternative in order to maximize another person’s suffering.
3. The decision maker asks his two children to rank the alternatives and then chooses the alternative that is the best on average.
4. The decision maker has an ideal point in mind and chooses the alternative that is closest to it.
5. The decision maker looks for the alternative that appears most often in the choice set.
6. The decision maker always selects the first alternative that comes to his attention.
7. The decision maker has an ordering in mind and always chooses the median element.

9 Utility Representation with Finite Choice Set

**Rubenstein Lecture 2*
MWG 3.C
Kreps 2.1
Jehle and Reny 1.1

Moving from Preferences to Utility Functions

Why not just stick with preference relations?

Now that we've developed a nice binary preference relation, why do we need utility functions?

- How would you maximize a preference relation? In other words, how do you choose the set $A \subset X$ that contains the best alternatives in the choice set X?
- It's more convenient to maximize a numerical function than a preference relation.

If *any* preference relation could be represented by a utility function, we could use utility functions instead of preference relations without loss of generality. In addition, utility theory gives the possibility of numerical representations carrying additional meanings, like a is preferred to b more than c is preferred to d.

Definition The function $U : X \to \mathbb{R}$ is a *utility function* representing the preference $\succsim$ if for all x and $y \in X$, $x \succsim y$ if and only if $U(x) \geq U(y)$.

Existence of a Utility Representation—Finite Choice Set

What assumptions are needed for a utility representation to exist?

The assumptions for the existence of a utility representation are straightforward: preferences need to be *complete* and *transitive*.[1] And we will start out with a finite choice set X.

[1] A "well-behaved" utility representation, however, will require several more assumptions.

Theorem

When the set X is finite, there is always a utility representation.

This proof will require a few more tools before we can execute it.

Tools for the proof

Definitions

Relation Properties

Suppose R is a relation on a set A.

1. Relation R is *reflexive* if xRx for every x $\in A$
2. Relation R is *symmetric* if xRy implies yRx for all x, y $\in A$.
3. Relation R is *transitive* if whenever xRy and yRz, then also xRz.

Equivalence Relation

A relation R on a set A is an *equivalence relation* if it is reflexive, symmetric and transitive.

Equivalence Class

Suppose R is an equivalence relation on a set A. Given any element $a \in A$, the *equivalence class* containing a is the subset $\{x \in A | xRa\}$ of A consisting of all the elements of A that relate to a. This set is denoted as $[a]$.

Partition

A *partition* on a set A is a set of subsets of A, such that the union of all the subsets equals A, and the intersection of any two different subsets is $\emptyset$.

Strict Preference

$x \succ y$ when both $x \succsim y$ and not $y \succsim x$.

Indifference

$x \sim y$ when $x \succsim y$ and $y \succsim x$.

Exercise U1.1

Show that $\sim$ is an equivalence relation. That is, it satisfies the following properties:

- *Reflexive*, $x \sim x$ for every $x \in X$.
- *Symmetric*, $x \sim y \longrightarrow y \sim x$.
- *Transitive*, if $x \sim y$ and $y \sim z$ then $x \sim z$.

Exercise U1.2

> Let $\succsim$ be a preference relation on a set X. Define $I(\mathrm{x})$ to be the set of all $\mathrm{y} \in X$ for which $\mathrm{y} \sim \mathrm{x}$. Show the set $\{I(\mathrm{x})|\mathrm{x} \in X\}$ is a *partition* of X

You can rely on the following theorem to answer Exercise 2:

Theorem: Suppose R is an equivalence relation on a set A. Then the set $\{[a]|a \in A\}$ of equivalence classes of R forms a partition of A.

One more tool for the proof

Definition: $a \in X$ is a *minimal element* if $x \succsim a$ for all $x \in X$.

Lemma: In any finite set $A \subseteq X$ there is a minimal element.

Proof: Let's use mathematical induction to prove (we'll need to rely on this proof for the "existence" proof)...

Step 1: If A is a singleton, then by completeness its only element has to be minimal.

Step 2: Show that if it is true for n, it is true for $n+1$.

- Let A have $n+1$ elements and let $x \in A$.
- The set $A - [x]$ has n elements.
- Assume $A - [x]$ has a minimal element. Show that A has a minimal element.
- (*The inductive hypothesis*) y is the minimal element of $A - [x]$.
- If $x \succsim y$ then y is minimal in A.
- If $y \succsim x$ then, by transitivity, $z \succsim x$ for all $z \in A - [x]$, therefore x is minimal.
- Q.E.D.

Utility Representation Proof (Finite X)

Claim

If $\succsim$ is a preference relation on a finite set X, then $\succsim$ has a utility representation with values being natural numbers.

Proof

Let X_1 be the subset of elements that are minimal in X, i.e., $X_1 = \{\text{x}|\text{x is minimal in } X\}$. Define

$$X_2 = \{\text{x}|\text{x is minimal in } X - X_1\} \tag{9.1}$$
$$\vdots \tag{9.2}$$
$$X_K = \{\text{x}|\text{x is minimal in } X - X_1 - X_2 - \cdots - X_{K-1}\} \tag{9.3}$$

Define $U(\text{x}) = 1$ if $\text{x} \in X_1$

Partition X.

1. $X = X_1 \cup X_2 \cup \cdots \cup X_k$
2. $X_{k+1} = \{\text{x}|\text{x is minimal in } X - X_1 - \cdots - X_k\}$
3. Continue until we have partitioned X, which we will eventually b/c X is finite.

Define $U(\text{x}) = k$ if $\text{x} \in X_k$.

To verify that U represents $\succsim$, let $\text{a} \succ \text{b}$. Then $\text{a} \notin X_1 \cup X_2 \cup \ldots \cup X_{U(\text{b})}$, thus $U(\text{a}) > U(\text{b})$.

Countably Infinite X

We can even show that there is a utility representation on an infinite choice set, X, as long as it is countable.

Claim

If X is countable, then any preference relation on X has a utility representation with range $(-1, 1)$.

Exercise U1.3

Prove this claim. NB: The proof is on page 14-15 of Rubinstein. Don't just copy it down. Show me that you understand it. Or give your own.

Lexicographic Preferences

Lexicographic preferences[2] are the outcome of applying the following procedure for determining the ranking of any two elements in the set X:

$$(\succsim_1, \succsim_2, \succsim_3, \cdots, \succsim_K)$$

These are different criteria applied sequentially. The lexicographic preferences induced by those preferences are x $\succsim_L$ y if

1. $\exists k^*$ such that for all $k < k^*$ we have x $\sim_k$ y and x $\succ_{k^*}$ y or
2. x $\sim_k$ y for all k.

Lexicographic preferences do not have a utility representation

Claim The lexicographic preference relation $\succsim_L$ on $[0,1]\times[0,1]$, induced from the relations x $\succsim_k$ y if $x_k \geq y_k$ $(k = 1, 2)$, does not have a utility representation.[3]

Reductio Proof Assume $u\colon X \to \mathbb{R}$ represents $\succsim_L$ where $X = [0,1] \times [0,1]$. For any $a \in [0,1]$ $(a,1) \succ_L (a,0)$. So, by definition of $u(\cdot)$, $u(a,1) > u(a,0)$.

Let $q(a)$ be a rational number in the nonempty interval I_a (here I stands for interval, not indifference). $I_a = (u(a,0), u(a,1))$.

1. $q(a)$ is a function from $[0,1]$ into the set of rational numbers
2. $q(a)$ is injective: if $b > a$ then $(b,0) \succ_L (a,1)$ which implies $u(b,0) > u(a,1)$. Define $I_b = (u(a,1), u(b,0))$
   a) I_a and I_b are disjoint
   b) $q(a) \neq q(b)$

Since u maps onto the real number line, the cardinality of the set of intervals equals the cardinality of $\mathbb{R}$. Since there exists a unique rational number in each interval, the set of rational numbers must have the same cardinality as $\mathbb{R}$, a contradiction. In other words q is injective and $q : \mathbb{R} \to \mathbb{Q}$, but $|\mathbb{R}| > |\mathbb{Q}|$, a contradiction.

[2] MWG talks about lexicographic preferences and gives a similar proof on page 46.
[3] Note that these preferences are not continuous because $(1,1) \succ (1,0)$ but in any ball around $(1,1)$ there are points inferior to $(1,0)$

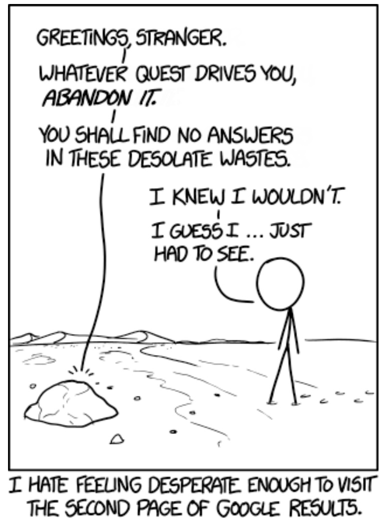

Figure 9.1: Second (xkcd.com)

10 Utility Representation with Infinite Choice Set

Rubenstein Lecture 2
MWG 3.C
Kreps 2.1
Jehle and Reny 1.1

Existence of a Utility Representation—Infinite Choice Set

What happens when X is infinite? What does it mean for X to be infinite? What are some examples of an *infinite* choice set?

Continuity of Preferences

If we're going to deal with an infinite choice set, X, we need to impose some "topological regularity" if we're going to have any success showing that there is a utility function that represents preferences over X. The "topological regularity" we will impose is the *continuity of preferences*.[1]

Defining continuous preferences

Continuous preferences, in a sense, means that preferences are "stable". The intuition is that if a is preferred to b, then small deviations from a or from b will not reverse the ordering.

Technical Note: We're going to start using a lot of ideas from topology. The first is an *epsilon ball* $B_\varepsilon(x)$. $B_\varepsilon(x)$ is a 'ball' around $x \in X$ with radius $\varepsilon > 0$. It is the set of all points in X that are distanced less than ε from x.

Definition 1: A preference relation $\succsim$ on X is *continuous* if whenever $a \succ b$ (namely it is not true that $b \succsim a$), there are balls $B_\varepsilon(a)$ and $B_\varepsilon(b)$ around a and b, respectively, such that for all $x \in B_\varepsilon(a)$ and $y \in B_\varepsilon(b)$, $x \succ y$.

[1] Here's where we start piling on the assumptions.

Definition 2 A preference relation $\succsim$ on X is *continuous* if the graph of $\succsim$ (i.e., the set $\{(x, y)|x \succsim y\} \subset X \times X$) is a closed set. That is, if $\{(a_n, b_n)\}$ is a sequence of pairs of elements in X satisfying $a_n \succsim b_n$ for all n and $a_n \to a$ and $b_n \to b$, then $a \succsim b$.[2]

Prove that Definition 1 and Definition 2 are Equivalent

Prove that the preference relation $\succsim$ on X satisfies the first definition of continuity if and only if it satisfies the second.

Purpose We're doing this to get some practice with the *reductio* proof method and to ease into some real analysis ideas.

Reminder

These two rules, LEM and modus tollens, enable the method of *proof by contradiction*, also know as proof by *reductio ad absurdum* or simply *reductio*. Here's how it works. Say you want to prove $P \to Q$. What you do is assume $P \to \neg Q$ and then show that $P \to \neg Q$ logically entails a contradiction like, say, 'Q'. By modus tollens, if $P \to \neg Q$ and $(Q \,\&\, \neg Q)$ is false (which it is by LEM), then $P \to \neg Q$ is false; and, by LEM, if $P \to \neg Q$ is false, then $P \to Q$ must be true.

First we will show that Definition 1 implies Definition 2, then vice versa.

Def 1 $\to$ Def 2

The first proposition we want to prove is Def 1 $\to$ Def 2. We will use a *reductio* proof by assuming Def 1 $\to$ not-Def 2 and showing that it creates a contradiction.

In particular, we make the following three assumptions:

1. Definition 1 holds.
2. The sequence (a_n, b_n) is a sequence of pairs satisfying $a_n \succsim b_n$ for all n and $a_n \to a$ and $b_n \to b$
3. (*The Reductio Assumption*) It is not true that $a \succsim b$. In other words, it is true that $b \succ a$.

[2] This definition is similar to MWG 3.C.1

P and $\neg Q$ These three assumptions amount to assuming that $\succsim$ on X is continuous according to Definition 1 and that Definition 2 does not hold.

Reductio Proof Based on Definition 1 and the assumption that $b \succ a$ there exist two balls $B_\varepsilon(a)$ and $B_\varepsilon(b)$ around a and b, respectively, such that for all $y \in B_\varepsilon(b)$ and $x \in B_\varepsilon(a)$, $y \succ x$.

Since $a_n \to a$ and $b_n \to b$, there is an N large enough that for all $n > N$ both $b_n \in B_\varepsilon(b)$ and $a_n \in B_\varepsilon(a)$.

That means that for some $n > N$ we have $b_n \succ a_n$.

But our second assumption was that $a_n \succsim b_n$, so we have a contradiction. □

$\therefore P \to Q$ Therefore if Definition 1 holds and $\{(a_n, b_n)\}$ is a sequence of pairs of elements in X satisfying $a_n \succsim b_n$ for all n and $a_n \to a$ and $b_n \to b$, then $a \succsim b$.

Def 2 → Def 1

Now, assume that $\succsim$ is continuous according to Definition 2 and show that it implies Definition 1.

Reductio Make the following three assumptions:

1. Definition 2 holds.
2. $a \succ b$.
3. (*The Reductio Assumption*) For all n there exists $a_n \in B_{1/n}(a)$ and $b_n \in B_{1/n}(b)$ such that $b_n \succsim a_n$.

The third assumption implies the sequence (b_n, a_n) converges to $(b, a) \forall n$ (how? be sure you understand this). By Definition 2, (b, a) is within the graph of $\succsim$, that is, $b \succsim a$. But this contradicts the second assumption. □

Useful Definitions

Convex Set A set C in n space is *convex* if it has the property that, whenever two points p and q are in C, then so are all points of the form $\lambda p + (1 - \lambda)q \quad 0 < \lambda < 1$.

Interior A point q is said to be *interior* to a set S if all the points sufficiently near to q also belong to S.

Interior of a Set The *interior of a set* S is the subset of S consisting of all points q that are interior to S.

Open A set S is called *open* if every point in S is interior.

Dense The set Y is dense in X if every open set $B \subset X$ contains an element in Y.

Utility Representation for Continuous Preferences

Proposition Assume that X is a convex subset of $\mathbb{R}^n$. If $\succsim$ is a continuous preference relation on X, then $\succsim$ has a utility representation.

Exercise U2.1

> Prove the proposition. (See Rubenstein p.19 if you get hung up, but don't just copy his proof. Show me you understand it, or give your own.)

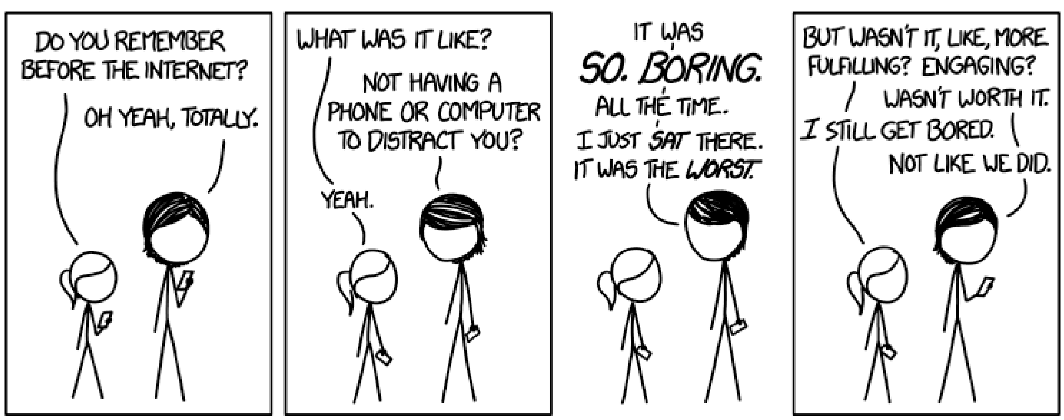

Figure 10.1: Before the internet (xkcd.com)

11 Preference Axioms

**Rubenstein Lecture 4*
MWG 3.A-C
Kreps 32-37
Jehle and Reny 1

The Choice Set for a Consumer

We are going to put some structure on the choice set, X.

- Restrict X to be $\mathbb{R}_+^K = \{\text{x} = (x_1, \ldots, x_K) |$ for all $k, x_k \geq 0\}$.
- An element of $\text{x} \in X$ is called a *bundle*.
- A bundle x is interpreted as a combination of K commodities, $\text{x} = (x_1, \ldots, x_K)$
- x_k is the quantity of commodity k

NB The minimal requirements on the consumption set are going to be:

1. $X \subseteq \mathbb{R}_+^k$
2. X is closed
3. X is convex
4. $0 \in X$

With the added structure of X we can use the properties of $\mathbb{R}^K$ to impose more structure on preferences. Here's a quick recap of what we have so far plus the new conditions.

The Preference Axioms

Completeness, Transitivity

Axiom 1 For all x and y in X, either $\text{x} \succsim \text{y}$ or $\text{y} \succsim \text{x}$ or both.

Axiom 2 For any three elements x, y and z in X, if x $\succsim$ y and y $\succsim$ z then x $\succsim$ z.

Example Let's say $X = \mathbb{R}^2$. Preferences that satisfy axioms 1 and 2 might look like this

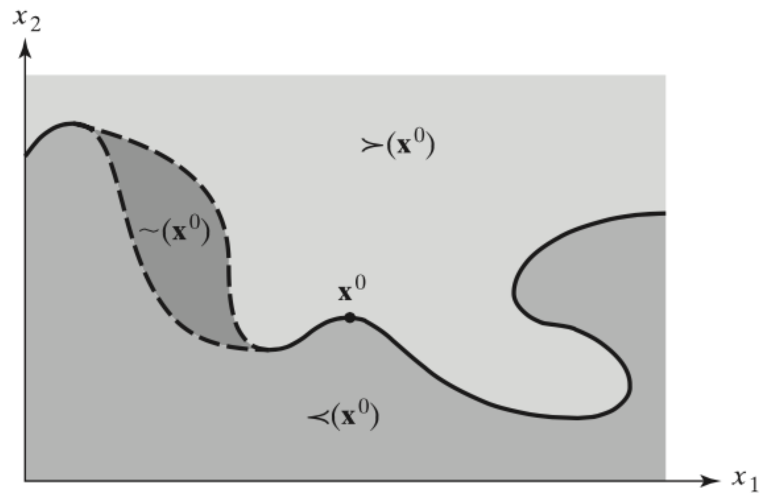

Figure 11.1: (J& R Figure 1.1)

Continuity

Continuity uses the topological structure of $\mathbb{R}^k$, which gives us the ability to talk about closeness.

Definition One more definition will be useful in the following proof. $\forall$ x,

1. $S^+(\mathrm{x}) = \{\mathrm{y}|\mathrm{y} \succsim \mathrm{x}\}$;
2. $S^-(\mathrm{x}) = \{\mathrm{y}|\mathrm{x} \succsim \mathrm{y}\}$.

In other words, $S^+(\mathrm{x})$ is the set of commodity bundles that are at least as good as x, while $S^-(\mathrm{x})$ is the set of commodity bundles that are no better than x.

Axiom 3 We can define the continuity of preferences in yet one more way.

$$\forall\, \mathrm{x},\ \text{both } S^+(\mathrm{x}) \text{ and } S^-(\mathrm{x}) \text{ are closed.}[1]$$

[1] A *closed* set is one such that if you take any converging infinite sequence of points in the set the limit is also in the set.

Exercise U3.1

> Do Rubenstein Problem Set 2, Problem 1.d, which asks you to show that this definition of continuity is equivalent to the two we saw in the Utility II lecture notes.

Axiom 3 eliminates sudden preference reversals.

Monotonicity

Monotonicity uses the ordering on the axis of $\mathbb{R}^k$, which gives us the ability to compare bundles by the amount of any particular commodity.

The relation $\succsim$ satisfies *monotonicity at the bundle* y if for all x $\in X$,

$$\text{if } x_k \geq y_k \text{ for all } k, \text{ then x} \succsim \text{y, and}$$

$$\text{if } x_k > y_k \text{ for all } k, \text{ then x} \succ \text{y.}$$

Axiom 4 The relation $\succsim$ satisfies *monotonicity* if it satisfies monotonicity at every y $\in X$.

This axiom says that "goods are goods." If the bundle x has at least as much of each commodity as the bundle y, and maybe more of some things, then the consumer likes x at least as much as y. Note that it is *not* required that the consumer prefer having more of any particular commodity, only that more of a something is not bad.

The relation $\succsim$ satisfies *strong monotonicity at the bundle* y if for all x $\in X$

$$\text{if } x_k \geq y_k \text{ for all } k \text{ and x} \neq \text{y, then x} \succ \text{y.}$$

Axiom $4'$ The relation $\succsim$ satisfies *strong monotonicity* if it satisfies strong monotonicity at every y $\in X$.

Axiom $4''$ Local Non-satiation

For all x $\in \mathbb{R}^n$, and for all $\epsilon > 0$, there exists some y $\in B_\epsilon(x) \cap \mathbb{R}^n$ such that y $\succ$ x.

This axiom states that the consumer is never wholly satisfied and is not even at a *relative* point of bliss. For any commodity bundle, x, and any tiny neighborhood of x, there is always some other commodity bundle in that neighborhood that the consumer likes better than x.

Note that local non-satiation rules out "zones of indifference", and it doesn't say anything about direction. Every preference relation that is monotonic at a bundle y is also nonsatiated at y, but the reverse is not true.

NB Monotonicity has a big impact on preference sets—they cannot 'bend upward'.

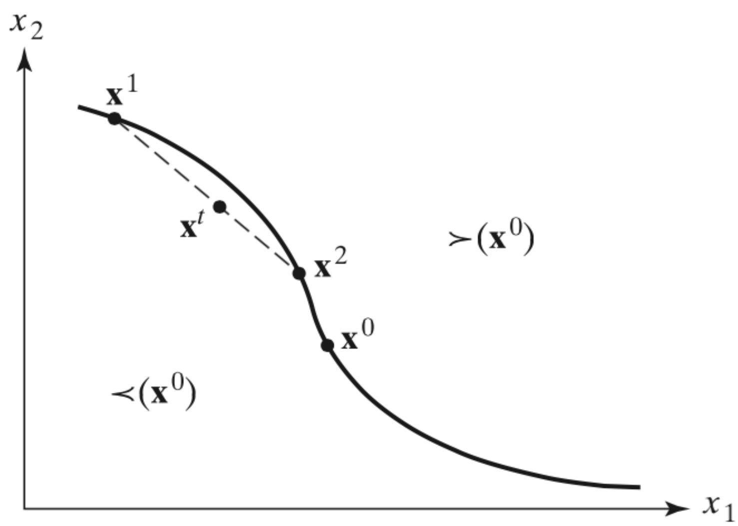

Figure 11.2: (J& R Figure 1.5)

In-class exercises

Suppose $X \subseteq \mathbb{R}^2$.

1. Show that the preference relation represented by $\min\{x_1, x_2\}$, i.e., x $\succsim$ y iff $\min\{x_1, x_2\} \geq \min\{y_1, y_2\}$, satisfies weak monotonicity.
2. Show that the preference relation represented by $\min\{x_1, x_2\}$ does not satisfy strong monotonicity.
3. Show that the preference relation represented by $x_1 + x_2$ satisfied strong monotonicity.
4. Show that the preference relation represented by the negative of the Euclidean distance, $-d(x, x^*) = -\sqrt{\sum(x_k - x_k^*)^2}$, is nonsatiated at every bundle except x^*.
5. Show that the preference relation represented by the negative of the Euclidean distance, $-d(x, x^*) = -\sqrt{\sum(x_k - x_k^*)^2}$, does not satisfy weak monotonicity.

The Consumer Wants More

Axioms 1-4 enable us to prove that there is always at least one commodity that the consumer would like to have in increased amounts (possibly a different one at different points). We show this by showing that the consumer would always like to have more of *everything*.

Lemma 1

$$\forall\, \mathrm{x}, \mathrm{z} \in X,\ \text{if } \mathrm{z} > \mathrm{x} \text{ then } \mathrm{z} \succ \mathrm{x}.$$

Proof

Since $\mathrm{z} > \mathrm{x}$, $\exists\, \delta > 0$ sufficiently small that $\forall\, \mathrm{y}$ so close to x that $\| \mathrm{y} - \mathrm{x} \| < \delta, \mathrm{z} > \mathrm{y}$ But by local non-satiation, $\exists$ at least one such y such that $\mathrm{y} \succ \mathrm{x}$. By weak monotonicity, however, $\mathrm{z} \succ \mathrm{y}$, and by the transitivity of preferences, $\mathrm{z} \succ \mathrm{x}$.

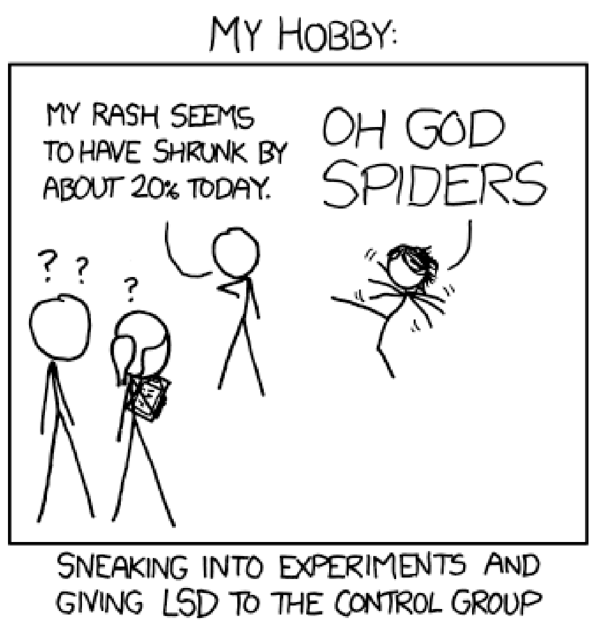

Figure 11.3: My hobby (xkcd.com)

12 The Utility Function

**Rubenstein Lecture 4*
MWG 3.A-C
Kreps 32-37
Jehle and Reny 1

Existence of the Utility Function

In the Utility II lecture, we demonstrated the existence of a utility function when X is a connected subset of $\mathbb{R}^n$. That proof did not rely on preferences satisfying monotonicity. Now that we have Axiom 4 we have a much simpler, more elegant proof. But first let me remind you of the definition of a utility function.

Definition The function $U : X \to \mathbb{R}$ is a *utility function* representing the preference $\succsim$ if for all x and $y \in X$, $x \succsim y$ if and only if $U(x) \geq U(y)$.

Theorem

Under Axioms 1-4, a utility function exists and is continuous in x.

Proof

First, the proof of existence. (Continuity is much harder and should not be attempted at home.) ;-)

Let e denote the n-vector e each of whose components in unity. Then for λ some positive scalar, λe is a point all of whose components are equal to λ.

Step 1 Note that, by Lemma 1, points in the form λe are strongly ordered. in other words, $\forall\, \lambda_1, \lambda_2$, if $\lambda_1 > \lambda_2$ then λ_1e $\succ$ λ_2e.

Step 2 Now, choose some x $\geqq 0$. By weak monotonicity, x $\succsim 0$. Choose $\overline{\lambda}$ larger than the greatest component of x. Then $\overline{\lambda}$e $>$ x, and, by Lemma 1, $\overline{\lambda}$e $\succ$ x.

Step 3 Consider points in the form λe, where $\lambda \in [0, \overline{\lambda}]$. As λ increases from 0 to $\overline{\lambda}$, we pass from points in S^-(x) to points in S^+(x). Further, because points with higher λ are preferred to points with lower ones, once we leave S^-(x), we can never reenter it.

Step 4 Hence there exists a special λ, say $\lambda(\text{x})$, with the property that $\forall\, \lambda < \lambda(\text{x})$, x $\succsim \lambda$e, while $\forall\, \lambda > \lambda(\text{x})$, λe $\succsim$ x. Take a sequence of λ's approaching $\lambda(\text{x})$ from below. For all λ in that sequence, $\lambda\text{e} \in S^{-}(\text{x})$. Hence, by the continuity of preferences, so is $\lambda(\text{x})$e. Similarly, take a sequence of λ's approaching $\lambda(\text{x})$ from above. For all λ in that sequence, $\lambda\text{e} \in S^{+}(\text{x})$. Hence, by the continuity of preferences, so is $\lambda(\text{x})$e. Thus x $\succsim \lambda(\text{x})$e, and also $\lambda(\text{x})\text{e} \succsim$ x. So $\lambda(\text{x})\text{e} \sim$ x.

Step 5 Define $u(\text{x}) = \lambda(\text{x})$. Is u(.) so defined a utility function? Well, take two commodity bundles, x and y, with x $\succsim$ y. By construction $\lambda(\text{x})\text{e} \sim$ x and $\lambda(\text{y})\text{e} \sim$ y. So, by Axiom 1, it follows that $\lambda(\text{x})\text{e} \succsim \lambda(\text{y})\text{e}$. But because points in the form λe are strongly ordered, this means $\lambda(\text{x}) \geq \lambda(\text{y})$.

On the other hand, suppose that $\lambda(\text{x}) \geq \lambda(\text{y})$. Then $\lambda(\text{x})\text{e} \succsim \lambda(\text{y})\text{e}$, by strong ordering. But x $\sim \lambda(\text{x})$e and y $\sim \lambda(\text{y})$e. So x $\succsim$ y.

Continuous Utility Function

Exercise U3.2

> Do Rubinstein Problem Set 4 problem 2. (There is a complicated proof in MWG 3.C.1, but you can come up with a much more straightforward proof. Hint: use epsilon balls)

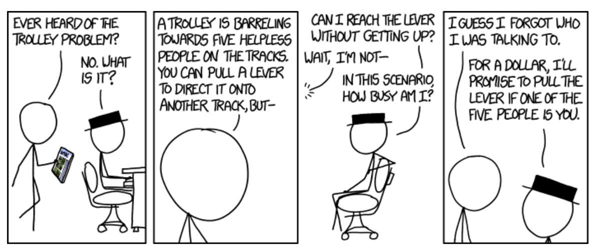

Figure 12.1: The trolley problem (xkcd.com)

13 Properties of Preferences and Utility Functions

**Rubenstein Lecture 4*
MWG 3.A-C
Kreps 32-37
Jehle and Reny 1

Convexity

We are going to impose convexity on the preference set. The "real world" interpretation is that consumers prefer a 'balanced' consumption basket.

Axiom 5 If $x \succsim y$, then $tx + (1-t)y \succsim y$ for all $t \in [0,1]$.

Axiom $5'$ Strict Convexity If $x \neq y$ and $x \succsim y$, then $tx + (1-t)y \succ y$ for all $t \in (0,1)$.

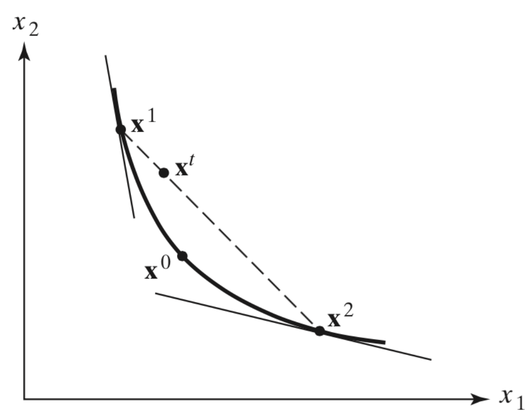

Figure 13.1: (J& R Figure 1.6)

Strict convexity requires the rate at which the consumer would trade y for x and remain indifferent to be strictly decreasing as we move from the north-west to the south-east along the indifference curve. This is called the *principle of diminishing marginal rate of substitution*.

The first two axioms, *completeness* and *transitivity*, impose "rationality," i.e., consumers are able to make consistent choices. The third axiom, *continuity*, is (mostly) for convenience—continuous preferences imply a continuous utility function. The last few axioms, *monotonicity* and *convexity*, impose structure on consumers' tastes, making our results going forward slightly less general than they otherwise would be. It's a small price to pay to simplify the proofs and discussion.

(Quasi)Concave Functions, Convex Combinations, and Convex Sets

Having placed (much) more structure on preferences—five axioms—we can now start talking about how this structure influences the utility function representation of preferences.

Before discussing some of the properties of preferences and utility functions induced by the axioms, let's be sure we're clear on some crucial definitions. These concepts are addressed more clearly on the following website:

http://www.economics.utoronto.ca/osborne/MathTutorial/QCC.HTM.

Definitions

Concave Function

$f : D \to \mathbb{R}$ is a *concave function* if for all $\mathrm{x}, \mathrm{y} \in D$,

$$f\left(\mathrm{tx} + (1 - \mathrm{t})\mathrm{y}\right) \geq t f(\mathrm{x}) + (1 - t) f(\mathrm{y}) \qquad \forall t \in [0, 1] \tag{13.1}$$

Examples

A function is concave if every straight line connecting two points on the surface lies everywhere on or under the surface.

Properties

Three important properties of concave functions:

1. their critical points are automatically global maxima
2. the weighted sum of concave functions is a concave function
3. the level sets of a concave function have just the right shapes for consumption and production theory.

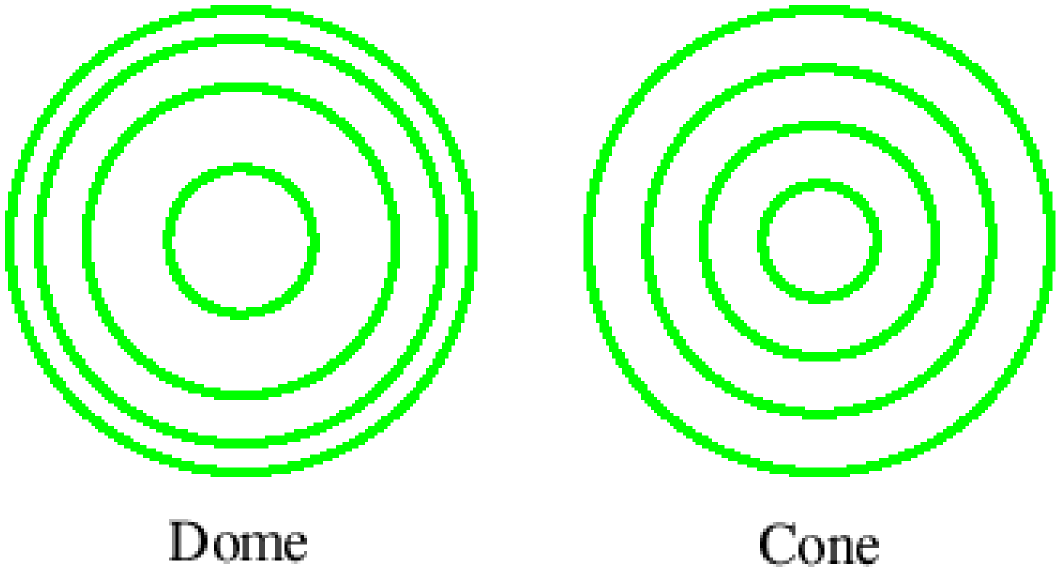

Figure 13.2: Level Sets

The properties of concave functions are largely *cardinal*, i.e., the distance between level sets means something. But in utility theory, we want functions that have *ordinal* meaning, not cardinal meaning. In other words, we are only concerned about the order of the level sets. It makes no sense to impose a stronger condition, like concavity, on the utility function, because the only significant property of the function is the character of its level curves, not the specific numbers assigned to these curves. So let's use a type of function that preserves the ordinal properties of concave functions without the cardinal properties.

Quasi-Concavity

A function u is *quasi-concave* if for all y the set $\{\mathrm{x} | u(\mathrm{x}) \geq u(\mathrm{y})\}$ is convex.

A function u defined on a convex set X is quasiconcave if every upper level set of u, $P_a = \{\mathrm{x} \in X | u(\mathrm{x}) \geq a\}$, is convex.[1]

Strict Quasiconcavity

A function $f : D \to \mathbb{R}$ is *strictly quasiconcave* iff, for all $\mathrm{x} \neq \mathrm{y}$ in D, $f(\mathrm{x}^t) > \min [f(\mathrm{x}), f(\mathrm{y})]$ for all $t \in (0, 1)$.

Every concave function is quasiconcave.

[1] Recall that a set A is convex if for all $a, b \in A$ and for all $\lambda \in [0, 1]$, $\lambda a + (1 - \lambda) b \in A$.

Quasiconcave functions lack the first two advantages of concave functions. For example, consider $y = x^3$. And what about $f_1(x) = x^3$ and $f_2(x) = -x$ and $f_3 = x^3 - x$?

Any monotonic transformation of a concave function is quasiconcave.

Preferences and Utility Functions

We have already seen that we can represent preferences with a continuous, real-valued function. A couple of theorems are in order:

Positive Monotonic Transforms

Theorem Let $\succsim$ be a preference relation on $\mathbb{R}^n$ and suppose $u(\mathrm{x})$ is a utility function that represents it. Then $v(\mathrm{x})$ also represents $\succsim$ if and only if $v(\mathrm{x}) = f(u(\mathrm{x}))$ for every x, where $f : \mathbb{R} \to \mathbb{R}$ is strictly increasing on the set of values taken on by u.

Exercise U4.1

Prove this.

Properties of Preferences and Utility Functions

Theorem Let $\succsim$ be represented by $u : \mathbb{R}^n \to \mathbb{R}$. Then:

1. $u(\mathrm{x})$ is strictly increasing if and only if $\succsim$ is strictly monotonic.
2. $u(\mathrm{x})$ is quasiconcave if and only if $\succsim$ is convex.
3. $u(\mathrm{x})$ is strictly quasiconcave if and only if $\succsim$ is strictly convex.

The Convexity Assumption

As we will see, assuming convex preferences will greatly simplify our ability to construct demand curves. But the assumption is not innocuous. When $X = \mathbb{R}^2$, the absolute value of the slope of the indifference curve is the *marginal rate of substitution of good two for good one*. It measures the rate at which the consumer is just willing to give up good two per unit of good one received.

Recall our ultimate goal is to have a theory of choice. We want to describe how a person chooses between two things (e.g., bundles). Soon we'll want to know how the comparison changes as prices change. But all the while we are comparing two things, and assumptions, like convexity, make it easier for us model the consumer's choice.

Marginal Utility

Before we look at budget constraints, we want to think about what it means to take the derivative of a utility function. The first-order partial derivative of $u(\mathrm{x})$ with respect to x_i is called the *marginal utility of good i*. We can derive an expression for the MRS using marginal utilities. Consider the bundle z $= (x, y)$. Because the indifference curve through z is just a function in the (x,y) plane, let $y = f(x)$ be the function describing it. Therefore, as x varies, the bundle $(x, y) = (x, f(x))$ traces out the indifference curve through z.

The MRS at bundle z, denoted

$$\mathrm{MRS}_{xy}(x, y) \equiv |f'(x)| = -f'(x) \tag{13.2}$$

because $f' < 0$. B/c $u(x, f(x))$ is constant, the derivative must be zero. That is,

$$\frac{\delta u(x,y)}{\delta x} + \frac{\delta u(x,y)}{y} f'(x) = 0 \tag{13.3}$$

Together, these imply that

$$\mathrm{MRS}_{12}(\mathrm{x}) = \frac{\frac{\delta u(\mathrm{x})}{\delta x}}{\frac{\delta u(\mathrm{x})}{\delta y}} \tag{13.4}$$

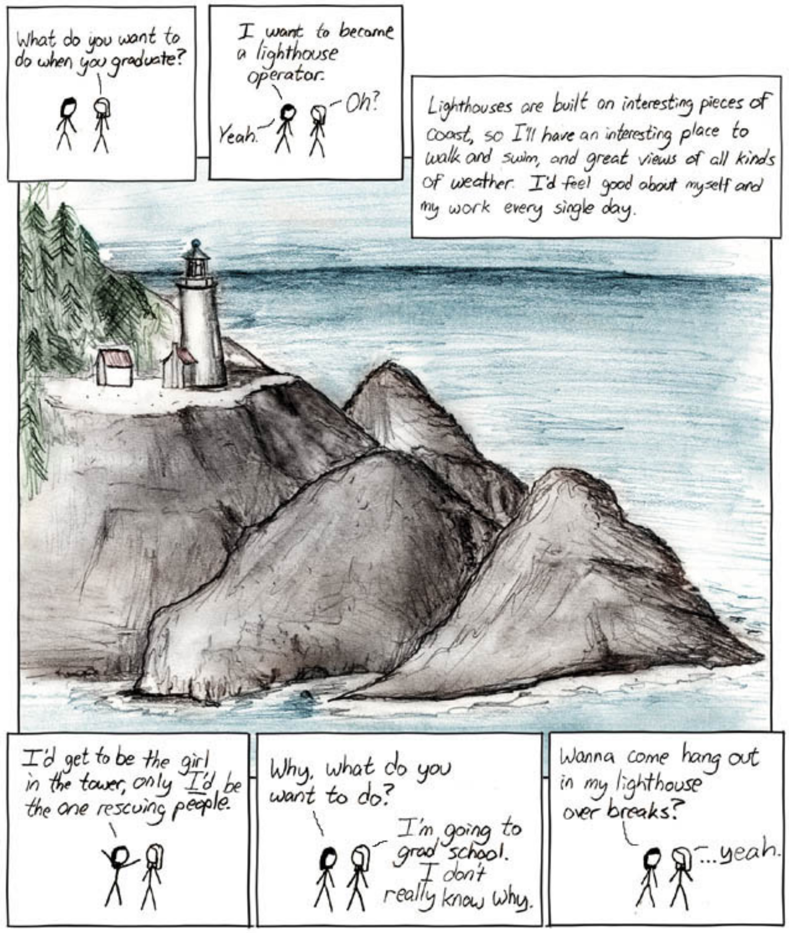

Figure 13.3: Lighthouse operator (xkcd.com)

14 The Consumers' Problem

**Jehle and Reny 1.3*
Rubenstein Lecture 5
Kreps p. 37-45
MWG 2, A–D, 3, D

Budget Set

So far we've talked about consumer preferences on an arbitrary set of bundles. But we're really interested in their choice on a particular type of set, called the budget set:

$$B(\mathrm{p}, w) = \{\mathrm{x} \in X | \mathrm{px} \leq w\} \tag{14.1}$$

where p is a vector of positive numbers we'll call prices, and w is a positive number we'll call wealth.

Properties of the budget set

$B(\mathrm{p}, w)$ is *compact*:[1]

- *closed*:[2] defined by weak inequalitites
- *bounded*:[3] for any $\mathrm{x} \in B(\mathrm{p}, w)$ and for all $k, 0 \leq x_k \leq \frac{w}{p_k}$

It is also *convex*:[4]

[1] *Heine-Borel Theorem*: The compact sets in n space are exactly those that are closed and bounded. The proof is left to you.

[2] The open ϵ-ball with center x_0 and radius ϵ>0 (a real number) is the subset of points in $\mathbb{R}^n$:

$$B_\epsilon(\mathrm{x}_0) \equiv \{\mathrm{x} \in \mathbb{R}^n | ||\mathrm{x}_0 - \mathrm{x}|| < \epsilon\} \tag{14.2}$$

The *closed* ϵ-ball with center x_0 and radius ϵ>0 (a real number) is the subset of points in $\mathbb{R}^n$:

$$B^*_\epsilon(\mathrm{x}_0) \equiv \{\mathrm{x} \in \mathbb{R}^n | ||\mathrm{x}_0 - \mathrm{x}|| \leq \epsilon\} \tag{14.3}$$

[3] A set S in $\mathbb{R}^n$ is called *bounded* if it is entirely contained within some ϵ-ball (either open or closed). That is, S is bounded if there exists some $\epsilon > 0$ such that $S \subset B_\epsilon(\mathrm{x})$ for some $\mathrm{x} \in \mathbb{R}^n$

[4] Let S be a vector space over the real numbers, or, more generally, some ordered field. A set $C \in S$ is said to be *convex* if, for all x and y in C and all α in the interval $[0, 1]$, the point $\alpha y + (1-\alpha)y$ is in C.

To show the budget set is convex, we need to show that any convex combination of two points in the budget set is also in the budget set.

- if $\mathrm{x}, \mathrm{y} \in B(\mathrm{p}, w)$ then
  - $\mathrm{px} \leq w$ and $\mathrm{py} \leq w$
  - $\mathrm{x}_k \geq 0$ and $\mathrm{y}_k \geq 0$.
- Thus, for all $\alpha \in [0, 1]$

$$\mathrm{p}(\alpha\mathrm{x} + (1-\alpha)\mathrm{y}) = \alpha\mathrm{px} + (1-\alpha)\mathrm{py} \leq w, \tag{14.4}$$

$$\alpha\mathrm{x}_k + (1-\alpha)\mathrm{y}_k \geq 0 \qquad \text{for all } k. \tag{14.5}$$

  that is, $\alpha\mathrm{x} + (1-\alpha)\mathrm{y} \in B(\mathrm{p}, w)$.

The Consumer Problem

Find the $\succsim$-best bundle in $B(\mathrm{p}, w)$.

Claim If $\succsim$ is a continuous relation,[5] then all consumer problems have a solution.

The proof of this general claim relies of some well-known theorems from real analysis. I will state the bottom line here, but you should look into how this is proven.

Theorem

If S is comp act, and f is continuous on S, then f takes a maximum and a minimum value somewhere on S. (The proof is in the appendix.)

Now we are ready to prove that all consumer problems have a solution as long as preferences, $\succsim$, are continuous.

Proof* If $\succsim$ is continuous, it can be represented by a continuous utility function u.

By the definition of "utility representation," finding an $\succsim$ optimal bundle is equivalent to solving the problem $\max_{\mathrm{x} \in B(\mathrm{p},w)} u(\mathrm{x})$.

Since the budget set is compact and u is continuous, the problem has a solution.

[5] $\succsim$ is continuous if for all $\mathrm{a}, \mathrm{b} \in X$, $\mathrm{a} \succ \mathrm{b}$ implies that there is an $\epsilon > 0$ such that $\mathrm{x} \succ \mathrm{y}$ for any x and y such that $||\mathrm{x} - \mathrm{a}|| < \epsilon$ and $||\mathrm{y} - \mathrm{b}|| < \epsilon$

Direct Proof

I know what you're thinking, "Gee whiz, that was fun, but I thought the utility function representation was just a convenience. It looks fairly fundamental in that proof. Can you do it without relying on the utility function?"[6]

Setup For any $\mathrm{x} \in B(\mathrm{p}, w)$ define the set $\textit{Inferior}(\mathrm{x}) = \{\mathrm{y} \in X | \mathrm{x} \succ \mathrm{y}\}$. By the continuity of preferences, every such set is open.

Proof This is a *reductio* proof

- Assume there is no solution to maximizing $\succsim$ on $B(\mathrm{p}, w)$.
- Then every $\mathrm{z} \in B(\mathrm{p}, w)$ is a member of some set $\textit{Inferior}(\mathrm{x})$. That is, the collection of sets $\{\textit{Inferior}(\mathrm{x}) | \mathrm{x} \in B(\mathrm{p}, w)\}$ covers $B(\mathrm{p}, w)$.[7]
- Recall that a collection of open sets that covers a compact set has a finite subset of sets that covers it.[8] Thus, there is a finite collection $\textit{Inferior}(\mathrm{x}_1), \ldots,$ $\textit{Inferior}(\mathrm{x}_n)$ that covers $B(\mathrm{p}, w)$.
- Letting x_j be the optimal bundle within the finite set $\{\mathrm{x}_1, \ldots, \mathrm{x}_n\}$, we obtain that $\mathrm{x}_j$is an optimal bundle in $B(\mathrm{p}, w)$, a contradiction.

Convexity and Solution Uniqueness

So, we've shown that the consumer problem has a solution. But how many solutions? And what do they look like? Where/how can we find them? Here is where assuming convex preferences has it's "bite."

Claim If $\succsim$ is convex, then the set of solutions for a choice from $B(\mathrm{p}, w)$ (or any other convex set) is convex.

[6]I know you're not *really* thinking this, because, seriously, who says 'gee whiz'?

[7]A collection $\mathcal{S}$ of open sets $\mathcal{O}_\alpha$ is said to be an *open covering* of the set S if $S \subset \bigcup \mathcal{O}_\alpha$. The covering is said to be a *definite covering* if $\mathcal{S}$ consists of only a finite number of open sets.

[8]A set S is called *compact* if every open covering of S can be reduced to a finite covering. This means that there must exist a finite sub-collection $\mathcal{S}_0$ of the original open sets which is still a covering of S.

Proof Assume that both x and y maximize $\succsim$ given $B(\mathbf{p}, w)$.

- By the convexity of the budget set B(p,w) we have $\alpha \mathrm{x} + (1-\alpha)\mathrm{y} \in B(\mathbf{p}, w)$.
- By the convexity of preferences, $\alpha \mathrm{x} + (1-\alpha)\mathrm{y} \succsim \mathrm{x} \succsim \mathrm{z}$ for all $\mathrm{z} \in B(\mathbf{p}, w)$.
- Thus $\alpha \mathrm{x} + (1-\alpha)\mathrm{y}$ is also a solution to the consumer problem.

Claim If $\succsim$ is strictly convex, then every consumer problem has at most one solution.

Proof Another *reductio* proof. Assume *both* x and y (where $\mathrm{x} \neq \mathrm{y}$) are solutions to the consumer problem $B(\mathbf{p}, w)$.

- Then $\mathrm{x} \sim \mathrm{y}$ (both are solutions to the same maximization) and $\alpha \mathrm{x}+(1-\alpha)\mathrm{y} \in B(\mathbf{p}, w)$ (the budget set is convex).
- By the strict convexity of $\succsim$, $\alpha \mathrm{x} + (1-\alpha)\mathrm{y} \succ \mathrm{x}$, which is a contradiction of x being a maximal bundle in $B(\mathbf{p}w)$.

Appendix

Definitions

Covering A collection $\mathcal{S}$ of open sets $\mathcal{O}_\alpha$ is said to be an *open covering* of the set S if $S \subset \bigcup \mathcal{O}_\alpha$. The covering is said to be a *definite covering* if $\mathcal{S}$ consists of only a finite number of open sets.

Compact A set S is called *compact* if every open covering of S can be reduced to a finite covering. This means that there must exist a finite sub-collection $\mathcal{S}_0$ of the original open sets which is still a covering of S.

Functions and Sets

Theorem Let f be a real-valued function defined and continuous on a compact set D. Then, f is bounded on D.

Proof

1. First, let's prove that any continuous function is locally bounded on its domain. Given $p_0 \in D$, take $\epsilon = 1$ and use the assumed continuity of f to choose a neighborhood U about p_0 so that $|f(p) - f(p_0)| < 1$ for all $p \in U \cap D$. Clearly, we than have $|f(p)| < 1 + |f(p_0)|$ for all such p, and f is locally bounded at p_0.
2. We will now use the compactness of D to obtain a single bound for $|f(p)|$ that works for all p in D. The fact that f is locally bounded everywhere in D, i.e., the result in 5, shows that given any $p \in D$, there is an open set $\mathcal{O}_p$ about p and a number M_p with $|f(q)| < M_p$ for all points $q \in \mathcal{O}_p$. The sets $\mathcal{O}_p$ form an open covering of D. Since D is compact, this covering can be reduced to a finite covering by sets $\mathcal{O}_{p_k}$ for $k = 1, 2, \ldots, m$. In each of these sets, f is bounded, and every point q in D lies in one of these sets. Thus, if M is the largest of the numbers M_{p_k}, we have $|f(q)| < M$ for all $q \in D$.

A real-valued function can be bounded and continuous on a set without attaining a maximum value on the set. This is true, for example, of the function $f(x) = x^2$ on the open interval $0 < x < 1$ and on the unbounded closed interval $0 \leq x \leq \infty$ for the function $f(x) = x/(1+x)$. The situation, however, is different on compact sets.

Theorem If S is compact, and f is continuous on S, then f takes a maximum and a minimum value somewhere on S.

Proof By the previous theorem, f is bounded on S. There then exist numbers B and b, with $b \leq f(p) \leq B$, for all $p \in S$. Let M be the smallest such upper bound for the values of f on S and m the largest lower bound. We must have $m \leq f(p) \leq M$ for all $p \in S$ as before, but now M cannot be decreased, nor m increased. If there is a point p_0 in S with $f(p_0) = M$, then M is the maximum value for f on S, and it is attained in S. If there is no point p_0 with $f(p_0) = M$, then $f(p) < M$ for all $p \in S$. In this case, set $g(p) = 1/(M - f(p))$. Since the denominator is never 0, g is continuous on S and must therefore be bounded there. If we suppose that $g(p) \leq A$ for all $p \in S$, then we would have

$$\frac{1}{M - f(p)} \leq A \quad \text{all}\, p \in S \tag{14.6}$$

and therefore

$$f(p) \leq M - \frac{1}{A} \tag{14.7}$$

This, however, contradicts the fact that M was the smallest upper bound for the values of f on S, and we conclude that this case does not occur and there is a point $p_0 \in S$ with $f(p_0) = M$. In a similar way we can show that there is a point in S where f has the value m, and a minimum is attained.

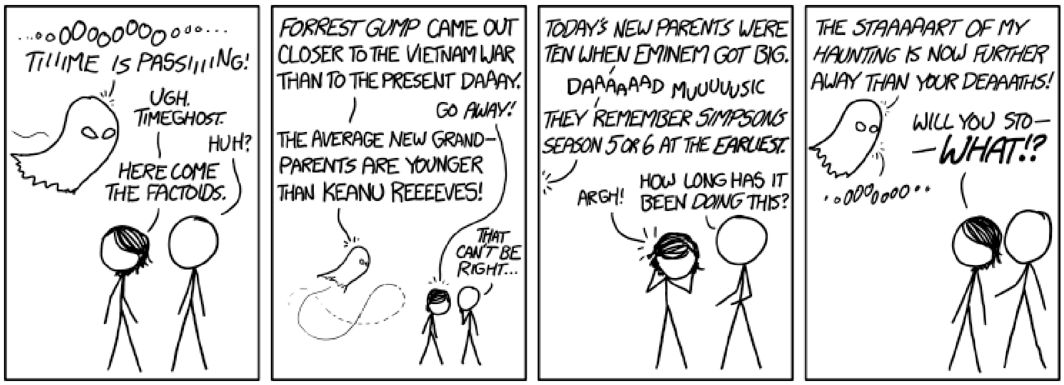

Figure 14.1: Time ghost (xkcd.com)

15 Utility Max

**Rubenstein Lecture 5*
Jehle and Reny 1.3
Kreps p. 37-45
MWG 2, A–D, 3, D

Bang-for-the-Buck: Necessary and Sufficient

As I've mentioned, I'm not exactly sure what it means for preferences to be differentiable. We will, however, presume that "marginal utility" means something that can be reflected in a slight change in preferences. Rubenstein calls this small preference change "subjective valuation." We'll think about it along those lines; that is, marginal utility is some sort of personal value of small changes in the consumption bundle. The objective, then, is to maximize the "bang for the buck"[1], i.e., the amount of personal value we get per dollar spent.

What follows establishes necessary and sufficient conditions for a solution to the consumers' problem (TCP). The necessary condition is that at a solution to TCP, the consumer equalizes the bang-for-the-buck across all commodities in the bundle.

Necessary

Assumptions: Assume that the consumers' preferences are differentiable with $\nu_i(\mathrm{x}^*) \quad i = 1, \ldots, K$ representing marginal utility of good i in bundle x^*. x^* is an optimal bundle in the consumer problem and k is a consumed commodity (i.e., $\mathrm{x}^*_k > 0$).

Claim: It must be that $\frac{\nu_k(\mathrm{x}^*)}{p_k} \geq \frac{\nu_j(\mathrm{x}^*)}{p_j}$ for all other j.

Reductio Proof: Assume that x^* is a solution to the consumer problem $B(\mathrm{p}, w)$ and that $\mathrm{x}^*_k > 0$ and, by way of contradiction, $\frac{\nu_j(\mathrm{x}^*)}{p_j} > \frac{\nu_k(\mathrm{x}^*)}{p_k}$.

A move in the direction of reducing the consumption of the k-th commodity by 1 and increasing the consumption of the j-th commodity by $\frac{p_k}{p_j}$ is an improvement since $\nu_j(\mathrm{x}^*)\frac{p_k}{p_j} - \nu_k(\mathrm{x}^*) > 0$.

[1] "Bang-for-the-buck" is an Americanism that means "value per dollar". You can, perhaps, think of "bang" as the impact, and "buck" is dollar.

Because $x^*_k > 0$, we can find $\epsilon > 0$ small enough such that decreasing k's quantity by ϵ and increasing j's quantity by $\epsilon\frac{p_k}{p_j}$ is feasible. This brings the consumer to a strictly better bundle (due to strong monotonicity), contradicting the assumption that x^*_k is a solution to the consumer problem. □

Implication If x^*_k is a solution to the consumer problem $B(p, w)$ and *both* $x^*_k > 0$ and $x^*_j > 0$, then the ratio $\frac{\nu_k(x^*)}{\nu_j(x^*)}$ must be *equal* to the price ratio $\frac{p_k}{p_j}$.

Sufficient

The sufficient condition is that if consumers have equalized the bang-for-the-buck across consumed commodities, then the consumed bundle is a solution to TCP. To establish this, we need convexity.

Assumptions $\succsim$ is strongly monotonic, convex, continuous, and differentiable, and at x^*

- $px^* = w$, and
- for all k such that $x^*_k > 0$, and for any commodity j, $\frac{\nu_k(x^*)}{p_k} \geq \frac{\nu_j(x^*)}{p_j}$.

Claim x^* is a solution to the consumer problem.

Reductio Proof If x^* is not a solution, then there is a bundle y such that $py \leq px$ and $y \succ x^*$. By continuity we can assume that $y_k > 0$ for all k.

Let $\mu = \frac{\nu_k(x^*)}{p_k}$ for all k with $x^*_k > 0$; consider μ as the bang-for-the-buck of consumed commodities. Because

1. y is feasible,
2. for a good k that is consumed, i.e., $x^*_k > 0$ we have $p_k = \nu_k(x^*)/\mu$, and
3. for a good k that is not consumed, i.e., $x^*_k = 0$, $(y_k - x^*_k) \geq 0$ and $\mu \geq \frac{\nu_k(x^*)}{p_k}$. In other words, if good k is not consumed, it's bang-for-the-buck must be (weakly) less than that of consumed goods. This can be rewritten as $p_k \geq \nu_k(x^*)/\mu$,

we can write:

$$0 \geq \mathrm{p}(\mathrm{y}-\mathrm{x}^*) = \sum p_k(\mathrm{y}_k - \mathrm{x}_k^*) \geq \sum \nu_k(\mathrm{x}^*)(\frac{\mathrm{y}_k - \mathrm{x}_k^*}{\mu}). \tag{15.1}$$

Thus, $0 \geq \nu(\mathrm{x}^*)(\mathrm{y}-\mathrm{x}^*)$, in contradiction to $(\mathrm{y}-\mathrm{x}^*)$ being an improvement direction.

Setup

Some assumptions

Assumption 1 Assume that $u(\cdot)$ is a continuous utility function that represents a weakly monotone, locally-nonsatiated preference relation $\succsim$.

Assumption 2 The preference relation $\succsim$ is strictly convex. That is, if $\mathrm{x} \neq \mathrm{y}$ and $\mathrm{x} \succsim \mathrm{y}$, then $t\mathrm{x} + (1-t)\mathrm{y} \succ \mathrm{y}$ for all $t \in (0,1)$.

Some notation

Let $X = \mathbb{R}^n_+$, $\mathrm{p} \in \mathbb{R}^n_+$.

Utility Maximization

$$v(\mathrm{p}, w) \equiv \max_{\mathrm{x}} u(\mathrm{x}) \text{ subject to } \mathrm{p} \cdot \mathrm{x} \leq w \text{ and } \mathrm{x} \geq 0 \tag{15.2}$$

- The solution to this program is $\mathrm{x}^*(\mathrm{p}, w)$: *Marshallian Demand*.
- v is the *indirect utility function*.

Figure 15.1: Integration by parts (xkcd.com)

16 Marshallian Demand

**Rubenstein Lecture 5*
Jehle and Philip 1.3
Kreps p. 37-45
MWG 2, A–D, 3, D

Consequences of Utility Maximization

First Characterization Theorem

What are the consequences of maximization? What restrictions on demand functions are implied by our assumption that demand is derived from the above utility maximization problem? Why is this important? Because we want to form *testable* hypothesis. We want to be able to empirically test whether our model of consumer behavior accurately explains *actual* behavior. To test the model it must impose restrictions on actual behavior. If consumer behavior obeys those restrictions, we can say it is *consistent with* the model of utility maximization.

Consider the utility maximization problem.

Characterization Theorem 1

1. If u is continuous, then for all (p, w) such that $\mathrm{p} \gg 0$ and $w > 0$, there exists at least one solution. If preferences are (weakly) convex, the set of optimizers is convex. If preferences are strictly convex, the solution is unique.

2. For all (p, w), $\mathrm{x}^*(\mathrm{p}, w) = \mathrm{x}^*(\alpha\mathrm{p}, \alpha\mathrm{w})$, $\alpha > 0$. That is, demand is homogeneous of degree 0.

Proof This follows from the basic equality $B(\alpha\mathrm{p}, \alpha w) = B(\mathrm{p}, w)$ and the assumption that the behavior of the consumer is "a choice from a set."

How well does this result reflect actual behavior?

3. If preferences are weakly monotone and locally insatiable, Assumption 1, then for all (p, w), $\mathrm{p} \cdot \mathrm{x}^*(\mathrm{p}, w) = w$. That is, any solution x^* to the consumer problem $B(\mathrm{p}, w)$ is located on its budget curve. This is referred to as *Walras' Law* or *Budget Balancedness*.

Proof If not, then $\mathrm{px} < w$. There is an $\epsilon > 0$ such that $\mathrm{p}(\mathrm{x}_1 + \epsilon, \ldots, \mathrm{x}_K + \epsilon) < w$. By monotonicity, $(\mathrm{x}_1 + \epsilon, \ldots, \mathrm{x}_K + \epsilon) \succ \mathrm{x}$, thus contradicting the assumption that x is optimal in $B(\mathrm{p}, w)$.

4. If $\succsim$ is a continuous preference, then the demand function is continuous in prices (and also in w).

Contrapositive Proof From Rubenstein Lecture 5.

Q Let $\succsim$ be a continuous preference relation.

By the continuity of preferences, there are neighborhoods B_{x^*}, B_{y^*} of x^* and y^* in which the strict preference is preserved.

P Show that the demand function is continuous in prices.

Not P Assume the demand function is not continuous in prices.

We want to show that $\neg P \rightarrow \neg Q$.

Show $\neg Q$ Let $\mathrm{x} \succ \mathrm{y}$. There are neighborhoods B_{x^*}, B_{y^*} of x^* and y^* in which the strict preference is *not* preserved, i.e., where $a \in B_{\mathrm{x}^*}$ and $b \in B_{\mathrm{y}^*}$ and $b \succ a$.

- By $\neg P$, there is a sequence of price vectors p_n converging to p^* such that $\mathrm{x}(\mathrm{p}^*, w) = \mathrm{x}^*$, and $\mathrm{x}(\mathrm{p}_n, w)$ does not converge to x^*.

  Thus we can assume that (p_n) is a sequence converging to p^* such that for all n the distance $||\mathrm{x}(\mathrm{p}_n, w) - \mathrm{x}^*|| > \epsilon$ for some positive ϵ.
- All numbers p_k are greater than some positive number m, so x_k is bounded by $0 \leq x_k \leq \frac{w}{m}$.
- By the *Bolzano-Weierstrass Theorem*[1] we can assume that $\mathrm{x}(\mathrm{p}_n, w) \rightarrow \mathrm{y}^*$ for some $\mathrm{y}^* \neq \mathrm{x}^*$.
- Since $\mathrm{p}_n\mathrm{x}(\mathrm{p}_n, w) \leq w$ for all n, it must be that $\mathrm{p}^*\mathrm{y}^* \leq w$, that is, $\mathrm{y}^* \in B(\mathrm{p}^*, w)$.
- Since x^* is the unique solution for $B(\mathrm{p}^*, w)$, we have $\mathrm{x}^* \succ \mathrm{y}^*$.
- For sufficiently large n, $\mathrm{x}(\mathrm{p}_n, w)$ is in B_{y^*}.

[1] B-W really has two parts. 1) Any bounded infinite set in n space has a cluster point. And 2) Any bounded sequence in n space has a limit point, and thus a converging subsequence.

- We want to find something from B_{x^*} that will contradict the continuity of preferences.
  We choose a bundle z^* in the neighborhood B_{x^*} such that $\mathrm{p}^*\mathrm{z}^* < w$.[2]
- Since $\mathrm{p}_n \to \mathrm{p}^*$, for all sufficiently large n, $\mathrm{p}_n\mathrm{z}^* < w$, which tells us that at price p_n, z^* is affordable (i.e., within the budget set).
- At price p_n, $\mathrm{x}(\mathrm{p}_n, w)$ is the optimal bundle,[3] which means $\mathrm{x}(\mathrm{p}_n, w) \succ \mathrm{z}^*$

$\neg Q$ In other words $\mathrm{x}^* \succ \mathrm{y}^*$ but $\mathrm{x}(\mathrm{p}_n, w) \succ \mathrm{z}^*$ where $\mathrm{x}(\mathrm{p}_n, w) \in B_{\mathrm{y}^*}$ and $\mathrm{z}^* \in B_{\mathrm{x}^*}$. So preferences are not continuous, a contradiction.

Adding Up Theorems

Budget balancedness imposes a structure on consumer behavior that we can see through the following adding up theorems.

Adding up theorem 1 A proportional change in all prices and income doesn't affect demand.

$$\text{For all } (\mathrm{p}, w), \text{ for all } i = 1, \ldots, n : \quad (\mathrm{p}, w) \cdot \nabla_{(\mathrm{p},w)} x_i^*(\mathrm{p}, w) = 0. \tag{16.1}$$

In other words,

$$\sum_{j=1}^{n} p_j \frac{\partial}{\partial p_j} x_i^*(\mathrm{p}, w) + w \frac{\partial}{\partial w} x_i^*(\mathrm{p}, w) = 0 \tag{16.2}$$

To prove this, use part (4) of the first characterization theorem and Euler's theorem.

Exercise 1 Let η represent the elasticity of demand for good i with respect to the price of good j. Let ε_i represent the elasticity of demand for good i with respect to income. Show that $\sum_{j=1}^{n} \eta_{ij} = -\varepsilon_i$. Interpret this result in words.

Adding up theorem 2 A change in the price of one good will not affect total expenditure.

$$\text{For all } (\mathrm{p}, w), \text{ for } i = 1, \ldots, n : \quad \sum_{j=1}^{n} p_j \frac{\partial}{\partial p_i} x_j^*(\mathrm{p}, w) + x_i^*(\mathrm{p}, w) = 0. \tag{16.3}$$

To prove this, use part (5) of the first characterization theorem.

[2] Since x^* is on the boundary, we want to make sure z^* is inside the budget set.
[3] By the definition of demand function.

Exercise 2 Let $\beta_i = \frac{p_i x_i^*(\mathrm{p},w)}{w}$, i.e., the proportion of income spent on good i. Show that $\sum_{j=1}^{n} \beta_j \eta_{ji} = -\beta_i$. Interpret in words.

Adding up theorem 3 A change in income will lead to a proportional change in total expenditure.

$$\sum_{j=1}^{n} p_j \frac{\partial}{\partial w} x_j^*(\mathbf{p}, w) = 1. \tag{16.4}$$

17 Indirect Preferences

**Rubenstein Lecture 6*
Jehle and Reny 1.3
Kreps p. 37-45
MWG 2, A–D, 3, D

Preferences Over Prices and Income

Typically, this is where the indirect utility function is introduced as a maximum-value function:

$$v(\mathbf{p}, w) = \max_{x \in \mathbb{R}^n} u(\mathbf{x}) \qquad s.t. \quad \mathbf{px} \leq w \tag{17.1}$$

This is purely a technical definition that allows us to employ a purely technical "trick"—the envelope theorem—to do comparative statics. Unfortunately, this approach doesn't really contain much economic intuition. So, we'll take a different approach that has more economic intuition, lays bare some of the limitations of the indirect approach, and leads to the same results.

Indirect Preference Relation

Let's think about the decision maker's preference relation over the domain of his primary choice function, D, e.g., budget sets. We can derive a preference relation over D, call it $\succsim^I$, by using the induced choice function from the preference relation $\succsim$ defined over X.

- A decision maker prefers set A over set B if the bundle he would choose from A is preferred to what he would choose from B.

$$A \succsim^I B \text{ if } C_{\succsim}(A) \succsim C_{\succsim} B \tag{17.2}$$

Call $\succsim^I$ the *indirect preference relation* induced from $\succsim$.

Representation of Indirect Preferences

If u represents $\succsim$ and the choice function is well defined, then $v(A) = u(C_{\succsim}(A))$ represents $\succsim^I$.

As with the direct approach, the indirect approach precludes certain reasonable-seeming behavior:

- I prefer $A - \{b\}$ to A even though I intend to choose a in any case since I am afraid to make a mistake by choosing b.
- I will choose a from A and from $A - \{b\}$; but since I don't want to have to reject b, I prefer $A - \{b\}$ to A.
- I prefer $A - \{b\}$ to A because I would choose b from A, and I want to commit myself to not making that choice.

Choosing Budget Sets

A consumer might choose budget sets, instead of bundles. For instance, one might considering the cost of living when deciding whether to move. Or think of a government comparing different tax systems that will affect p and w. If $\succsim$ satisfies all 5 axioms, we can define the indirect preferences $\succsim^I$ on the set $\mathbb{R}^{K+1}_{++}$ (i.e., the set of budget sets) as:

$$(\mathrm{p}, w) \succsim^I (\mathrm{p}', w') \text{ if } \mathrm{x}(\mathrm{p}, w) \succsim \mathrm{x}(\mathrm{p}', w'). \tag{17.3}$$

Properties of Indirect Preferences

1. *Invariance to presentation*: $(\lambda \mathrm{p}, \lambda w) \sim^I (\mathrm{p}, w)$ for all $\mathrm{p}, w, \lambda > 0$.
2. *Monotonicity*: Weakly decreasing in p_k and strictly increasing in w.
3. *Continuity*: If $(\mathrm{p}, w) \succ^I (\mathrm{p}', w')$, then $\mathrm{y} = \mathrm{x}(\mathrm{p}, w) \succ \mathrm{x}(\mathrm{p}', w') = \mathrm{y}'$.
4. *"Concavity"*: If $(\mathrm{p}_1, w_1) \succsim^I (\mathrm{p}_2, w_2)$, then $(\mathrm{p}_1, w_1) \succsim^I (\lambda \mathrm{p}_1 + (1-\lambda)\mathrm{p}_2, \lambda w_1 + (1-\lambda) w_2)$ for all $1 \geq \lambda \geq 0$.

Exercise IP.1

> Use the properties of *direct* preferences to sketch a proof of (6) and (7), and formally prove (1). The answers can be found in Rubinstein Lecture 6, but please try it on your own before checking Rubinstein.

A sketched proof of (2): Any bundle in $B\left(\lambda \mathrm{p}_1 + (1-\lambda)\mathrm{p}_2, \lambda w_1, (1-\lambda) w_2\right)$ is also in $B(\mathrm{p}_1, w_1)$ or $B(\mathrm{p}_2, w_2)$, so it's no better than the most preferred bundle.

A direct proof of (2):

1. Let z be the $\succsim$-best bundle in the budget set $B\left(\lambda \mathrm{p}_1 + (1-\lambda)\mathrm{p}_2, \lambda w_1 + (1-\lambda) w_2\right) = B(\mathrm{p}^\lambda, w^\lambda)$.

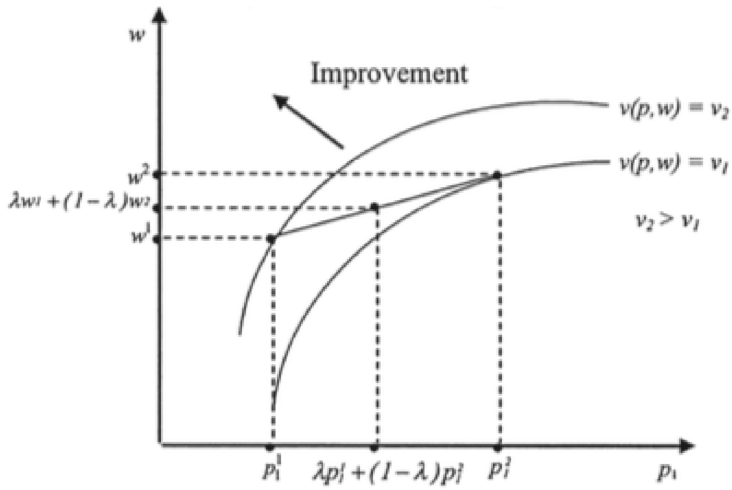

Figure 17.1: Rubinstein Figure 6.1

2. So, by budget balancedness, $p^\lambda z \leq w^\lambda$.
3. Therefore $p_1 z \leq w_1$ or $p_2 z \leq w_2$ (draw the three budget sets and convince yourself this is true). So $z \in B(p_1, w_1)$ or $z \in B(p_2, w_2)$
4. If I can spend all of my money and purchase exactly the same bundle, z, at some other price and income then the optimal bundle at the new price and income will be at least as good as z. Thus, $x(p_1, w_1) \succsim z$ or $x(p_2, w_2) \succsim z$.
5. From the premise in property 2 and equation (??), $x(p_1, w_1) \succsim x(p_2, w_2)$. It follows that $x(p_1, w_1) \succsim z$, which implies $(p_1, w_1) \succsim^I (\lambda p_1 + (1-\lambda)p_2, \lambda w_1 + (1-\lambda)w_2)$. This establishes the "concavity" of indirect preferences as illustrated in figure 17.1.

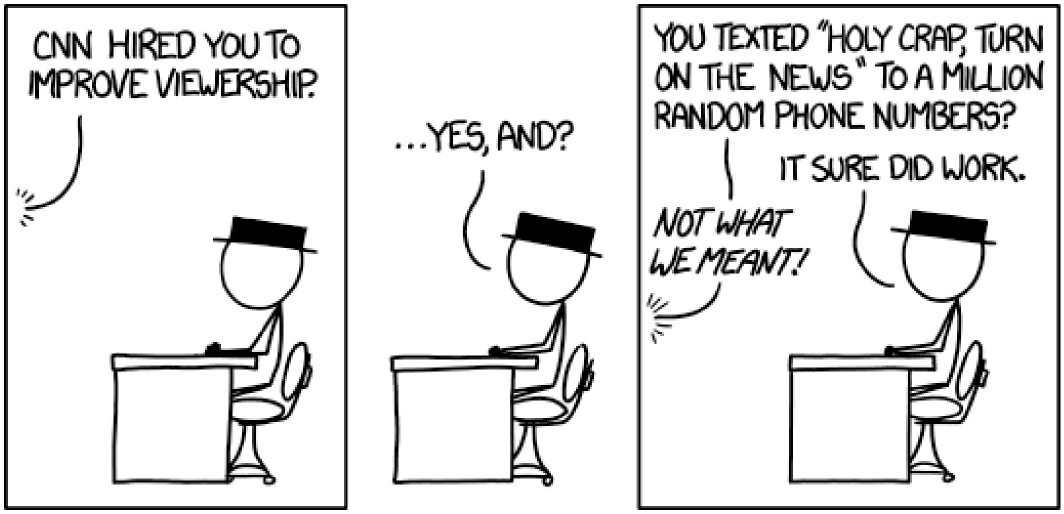

Figure 17.2: Mobile marketing (xkcd.com)

18 Marshalian Demand II

**Rubenstein Lecture 6*
Jehle and Reny 1.3
Kreps p. 37-45
MWG 2, A–D, 3, D

What can we do with indirect preferences? Can we get the demand function from indirect preferences? The answer is yes.

Roy's Identity

First, consider the situation with a single commodity. In this situation, the $\succsim^*$-indifference curve is a ray with slope w/p_1 through (p_1, w). This slope, of course, equals demand for the commodity.

More generally, if we have the slope of the indirect preferences indifference curve (ipic) through (p^*, w^*), we can recover demand at (p^*, w^*). They key is that the set $\{(\mathrm{p}, w) | \mathrm{px}(\mathrm{p}^*, w^*) = w\}$ is tangent to the ipic through (p^*, w^*).

We can derive demand from indirect preferences—if they are represented by a continuous, differentiable function $v(\mathrm{p}, w)$—using something called *Roy's Identity*:

$$x_k(\mathrm{p}^*, w^*) = -\frac{\frac{\partial v}{\partial p_k}(\mathrm{p}^*, w^*)}{\frac{\partial v}{\partial w}(\mathrm{p}^*, w^*)}. \tag{18.1}$$

Roy's great insight was that the vector of partial derivatives,

$$\left(\frac{\partial v}{\partial p_1}(\mathrm{p}^*, w^*), \ldots, \frac{\partial v}{\partial p_K}(\mathrm{p}^*, w^*), \frac{\partial v}{\partial w}(\mathrm{p}^*, w^*)\right) \tag{18.2}$$

is proportional to the (augmented) vector of of demands,

$$(x_1(\mathrm{p}^*, w^*), \ldots, x_K(\mathrm{p}^*, w^*), -1)\,. \tag{18.3}$$

Deriving Roy's Identity in Two Parts

Let's put ourselves in Roy's shoes and think about how we would come up with equation (18.1). In other words, how can we connect vector (18.2) with vector (18.3)?

The answer, essentially, is that we want to come up with two "lines" in $(\mathbf{p}, w)$ space: one will be a function of the relevant partial derivatives, and the other will be a function of demand. We'll show that these two "lines" coincide, and in that way we'll relate demand to (the partial derivatives of) the indirect utility function.

Hyperplanes

Below, we'll need to use hyperplanes.

But first, consider the level curve of the indirect utiilty function in (p, w) space. (Illustration provided in class.) The point (p_0, w_0) is on the indirect preference indifference curve (ipic). Below we will need a general way to define the line tangent to the ipic at (p_0, w_0). Let $\mathbf{r}_0$ be the position vector for (p_0, w_0). Now take a look at an arbitrary point on the tangent line: (p_1, w_1). Call $\mathbf{r}$ the position vector for (p_1, w_1). No matter where $\mathbf{r}$ is, the vector $\mathbf{r} - \mathbf{r}_0$ will always be in the tangent line, so we will use $\mathbf{r} - \mathbf{r}_0$ to characterize the tangent line more generally.

Now to hyperplanes:

- A *hyperplane* is a generalization of the plane into a different number of dimensions. A hyperplane of an n-dimensional space is a flat subset with dimension n−1
- *Supporting hyperplane* is a concept in geometry. A hyperplane divides a space into two half-spaces. A hyperplane is said to support a set in Euclidean space if the hyperplane meets both of the following conditions:
  - it is entirely contained in one of the two closed half-spaces determined by the hyperplane
  - it has at least one point on the hyperplane.

  Here, a closed half-space is the half-space that includes the hyperplane.

Part 1

As you surely remember from vector calculus, vector (18.2) is the gradient of $v(\mathbf{p}, w)$, the indirect utility function. This gradient is orthoganal to the level set of v.[1] Thus,

[1]If your memory is a little foggy on this, grab a multivariable calculus book to remind yourself this is true.

a hyperplane that is orthogonal to vector (18.2) at (p^*, w^*) will be tangent to the indirect preferences indifference curve (ipic).[2] In other words, the hyperplane

$$T = \left\{ (\mathrm{p}, w) \middle| \left(\frac{\partial v}{\partial p_1}(\mathrm{p}^*, w^*), \ldots, \frac{\partial v}{\partial p_K}(\mathrm{p}^*, w^*), \frac{\partial v}{\partial w}(\mathrm{p}^*, w^*) \right) \cdot (\mathrm{p} - \mathrm{p}^*, w - w^*) = 0 \right\} \tag{18.4}$$

is the unique tangent to the ipic through (p^*, w^*)[3].

Part 2

We can establish proportionality, hence Roy's identity, if we can find another hyperplane that is similarly tangent to the ipic and perpendicular to vector (18.3).

"The key observation," as Rubenstein says, "is that the set

$$H = \{(\mathrm{p}, w) \mid \mathrm{p} \cdot \mathrm{x}(\mathrm{p}^*, w^*) = w\} \tag{18.5}$$

is tangent to the ipic through (p^*, w^*)." This is much less obvious than Part 1, and presumably why Roy gets memorialized.

Graphical Intuition

First, note that $\mathrm{p} \cdot \mathrm{x}(\mathrm{p}^*, w^*) = w$ can be written as a hyperplane, H, in $p_1 \times w$ space,

$$p_1 x_1^* + \ldots + p_K x_K^* = w, \tag{18.6}$$

where the slope of the line is x_1^* and the intercept is $p_2 x_2^* + \ldots p_K x_K^*$.

To establish tangency, we need to show that every point in the hyperplane is "above" the ipic, except for (p^*, w^*).

Let's break this down.

- The fundamental idea is that starting from a budget set $B(\mathrm{p}^*, w^*)$, any $B(\mathrm{p}, w)$ for which $\mathrm{p} \cdot \mathrm{x}(\mathrm{p}^*, w^*) = w$ cannot be inferior to $B(\mathrm{p}^*, w^*)$.

  $\mathrm{x}(\mathrm{p}^*, w^*)$ is the $\succsim$-best bundle under prices p^* and income w^*. If I can spend all of my money and still be able to purchase the same bundle $(\mathrm{x}(\mathrm{p}^*, w^*))$ at some other set of prices and income, the $\succsim$-best bundle at the new (p, w) will be at least as good as $\mathrm{x}(\mathrm{p}^*, w^*)$.

[2] Remember that two vectors are orthogonal if their dot product is zero.
[3] Assume positive consumption of every good.

- In other words, for any $(\mathrm{p}, w) \in H$, the bundle $\mathrm{x}(\mathrm{p}^*, w^*) \in B(\mathrm{p}, w)$, and $\mathrm{x}(\mathrm{p}, w) \succsim \mathrm{x}(\mathrm{p}^*, w^*)$.
- Therefore, by the definition of $\succsim^I$, $(\mathrm{p}, w) \succsim^I (\mathrm{p}^*, w^*)$.[4]

That is, all points in the hyperplane H are "above" the ipic except (p^*, w^*), which is exactly on the ipic.

Back to the proof

Here, the great insight relies on the idea that indirect preferences are preferences for budget sets, and every point in $\mathrm{p} \times w$ space represents a particular budget set. So a hyperplane that coincides with T needs to be a linear combination of budget sets, including the one that resulted in $\mathrm{x}(\mathrm{p}^*, w^*)$. In other words, the line needs to be $\mathrm{p} \cdot \mathrm{x}(\mathrm{p}^*, w^*) = w$.

Rewrite equation (18.4) as $H = \{(\mathrm{p}, w) \mid (\mathrm{x}(\mathrm{p}^*, w^*), -1) \cdot (\mathrm{p}, w) = 0\}$. Combine this with the fact that $w^* = \mathrm{p}^* \cdot \mathrm{x}(\mathrm{p}^*, w^*)$ and we have:

$$H = \{(\mathrm{p}, w) \mid (\mathrm{x}(\mathrm{p}^*, w^*), -1) \cdot (\mathrm{p} - \mathrm{p}^*, w - w^*) = 0\}\,. \tag{18.7}$$

So $T = H$ and, H and T are both tangent to the ipic at (p^*, w^*). We conclude that vector (18.2) is proportional to the vector (18.3). Since vector (18.3) has -1 as its final element, we can easily deduce the coefficient of proportionality: $\frac{-1}{\frac{\partial v}{\partial w}(\mathrm{p}^*, w^*)}$.

Recap

So, what have we learned?

1. We can think of the "indirect" choice problem as choosing among budget sets using "indirect preferences," but indirect preferences are even more restrictive than direct preferences.
2. We can think of Roy's identity as a type of "marginal rate of substitution" between prices and income. In other words, the tradeoff between the price of good k and income necessary to maintain maximum utility, is the (negative of) the quantity being purchased.

[4] The line tangent to the ipic represents all of the budget sets that are at least as preferred as (p^*, w^*).

3. Purchasing $\mathrm{x}(\mathrm{p}^*, w^*)$ must always be an option when approaching the problem this way—this is the "constraint."

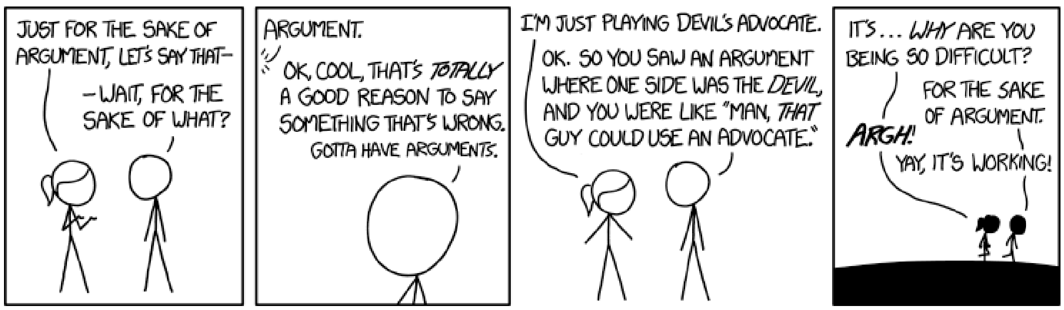

Figure 18.1: For the sake of argument (xkcd.com)

19 Compensated Demand

Jehle and Philip p33-41
**Rubenstein Lecture 6*
Kreps p51-59
MWG 3D-G

Expenditure Minimization

Traditional Definition

$$e(\mathrm{p}, \bar{u}) \equiv \min_{\mathrm{x}} \mathrm{p} \cdot \mathrm{x} \text{ subject to } u(\mathrm{x}) = \bar{u} \text{ and } \mathrm{x} \geq 0. \tag{19.1}$$

- The solution to this program is $\mathrm{x}^{**}(\mathrm{p}, \bar{u}) = \mathrm{h}(\mathrm{p}, \bar{u})$: *Hicksian* or *compensated* demand.[1]
- e is the *expenditure function*.

In Terms of Preferences

Another way of thinking about the EMP is to consider a special type of consumer who has in mind a bundle z and wishes to consume the cheapest bundle that is at least as good as z.

$$D(\mathrm{p}, \mathrm{z}) = \min_{\mathrm{x}} \{\mathrm{px} \big| \mathrm{x} \succsim \mathrm{z}\} = \min_{\mathrm{x}} \{\mathrm{px} \big| u(\mathrm{x}) \geq u(\mathrm{z}) = \bar{u}\}. \tag{19.2}$$

This is the *dual problem*. The solution is the *Hicksian demand function*, $h(\mathrm{p}, \bar{u})$. The expenditure function is $e(\mathrm{p}, \bar{u}) = \mathrm{p}h(\mathrm{p}, \bar{u})$.

Properties of the Expenditure Function

On the flip side, we might consider minimizing total expenditure to achieve a set utility level. Why would we want to do this when we already have a perfectly good way of estimating demand (utility max)? To motivate the issue, consider a town that wants to decrease the price of transportation for its residents by implementing a mass transit system. The town will have to raise taxes to pay for the mass transit system,

[1] x^* and h are vectors, so demand for good i is x_i^* or h_i

and they want to know whether its residents ultimately will be better (no worse) off. We might think of this as an expenditure-minimization-type question; the town will change relative prices and income to leave the residents at the same level of utility.

If u is continuous and strictly increasing then $e(\mathbf{p}, u)$ is

1. Zero at $u(\mathbf{0})$,
2. Continuous on its domain,
3. For all $\mathbf{p} \gg 0$, strictly increasing and unbounded above in u,
4. Increasing in $\mathbf{p}$,
5. Homogeneous of degree one in $\mathbf{p}$,
6. Concave in $\mathbf{p}$,
7. And if u is strictly quasiconcave, $e(\mathbf{p}, u)$ is differentiable in $\mathbf{p}$ at $(\mathbf{p}^0, u^0)$ whenever $\mathbf{p}^0 \gg 0$, and

$$\frac{\partial e(\mathbf{p}^0, u^0)}{\partial p_i} = h_i(\mathbf{p}^0, u^0), \qquad i = 1, \ldots, n. \tag{19.3}$$

Proof of 3 By way of contradiction, suppose that $e(\mathbf{p}, u)$ is not strictly increasing in u.

- let $\mathbf{x}'$ and $\mathbf{x}''$ denote optimal consumption bundle for required utility levels u' and u'', respectively, where $u'' > u'$ and $\mathbf{p} \cdot \mathbf{x}' \geq \mathbf{p} \cdot \mathbf{x}'' > 0$.
- Consider a consumption bundle $\tilde{\mathbf{x}} = \alpha \mathbf{x}''$, where $\alpha \in (0, 1)$.
- By continuity of $u(\cdot)$, there exists an α close enough to 1 such that $u(\tilde{\mathbf{x}}) > u'$ and $\mathbf{p} \cdot \mathbf{x}' > \mathbf{p} \cdot \tilde{\mathbf{x}}$.
- But this contradicts $\mathbf{x}'$ being optimal in the EMP with required utility level u'.

Proof of 4 To show that $e(\mathbf{p}, u)$ is nondecreasing in p_l:

- Suppose that the input price vectors $\mathbf{p}''$ and $\mathbf{p}'$ have $\mathbf{p}''_l \geq \mathbf{p}'_l$ and $\mathbf{p}''_k = \mathbf{p}'_k$ for all $k \neq l$.
- Let $\mathbf{x}''$ be an optimizing vector in the EMP for prices $\mathbf{p}''$.
- Then $\mathrm{e}(\mathbf{p}'', u) = \mathbf{p}'' \cdot \mathbf{x}'' \geq \mathbf{p}' \cdot \mathbf{x}'' \geq e(\mathbf{p}', u)$, where the latter inequality follows from the definition of $e(\mathbf{p}', u)$.

Proof of 5 The constraint set of the EMP is unchanged when prices change. Thus, for any scalar $\alpha > 0$, minimizing $(\alpha \mathbf{p}) \cdot \mathbf{x}$ on this set leads to the same optimal consumtion bundles as minimizing $\mathbf{p} \cdot \mathbf{x}$. Letting $\mathbf{x}^*$ be optimal in both circumstances, we have $e(\alpha \mathbf{p}, u) = \alpha \mathbf{p} \cdot \mathbf{x}^* = \alpha e(\mathbf{p}, u)$.

Proof of 6 For concavity, fix a required utility level $\bar{u}$, and let $\mathbf{p}'' = \alpha \mathbf{p} + (1-\alpha)\mathbf{p}'$ for $\alpha \in [0, 1]$. Suppose that $\mathbf{x}''$ is an optimal consumption bundle in the EMP when prices are $\mathbf{p}''$. If so,

$$\begin{aligned} e(\mathbf{p}'', \bar{u}) &= \mathbf{p}'' \cdot \mathbf{x}'' \\ &= \alpha \mathbf{p} \cdot \mathbf{x}'' + (1-\alpha)\mathbf{p}' \cdot \mathbf{x}'' \\ &\geq \alpha e(\mathbf{p}, \bar{u}) + (1-\alpha) e(\mathbf{p}', \bar{u}). \end{aligned} \tag{19.4}$$

where the last inequality follows because $u(\mathbf{x}'') \geq \bar{u}$ and the definition of the expenditure function implies that $\mathbf{p} \cdot \mathbf{x}'' \geq e(\mathbf{p}, \bar{y})$ and $\mathbf{p}' \cdot \mathbf{x}'' \geq e(\mathbf{p}', \bar{u})$.

Compensated/Hicksian Demand

Negative Own-Substitution Terms

Let $h_i(\mathbf{p}, \bar{u})$ be the Hicksian demand for good i. Then

$$\frac{\partial h_i(\mathbf{p}, \bar{u})}{\partial p_i} \leq 0, \qquad i = 1, \ldots, n. \tag{19.5}$$

Equation (??) tells us something about the way the unobservable Hicksian demand for a good changes when its own price changes. We will see that this gives us qualitative information, or *sign restrictions*, on the way the system of demand functions responds to price changes. But what if the price of another good changes?

Symmetric Substitution Terms

Let $\mathbf{h}(\mathbf{p}, \bar{u})$ be the consumer's system of Hicksian demands and suppose that $e(\cdot)$ is twice continuously differentiable. Then

$$\frac{\partial h_i(\mathbf{p}, \bar{u})}{\partial p_j} = \frac{\partial h_j(\mathbf{p}, \bar{u})}{\partial p_i}, \qquad i, j = 1, \ldots, n. \tag{19.6}$$

According to Jehle and Reny 2001, "One is very hard pressed to directly comprehend the significance of this result. Remarkably, however, it can be shown that this symmetry condition is intimately related to the assumed transitivity of the consumer's preference relation![sic]"

The preceding two theorems tell us a lot about the entire system of Hicksian substitution terms. We will use these properties of Hicksian demand to derive testable hypotheses.

The Substitution/Slutsky Matrix

Let $\mathrm{h}(\mathrm{p}, \bar{u})$ be the consumer's system of Hicksian demands, and let

$$\sigma(\mathrm{p}, \bar{u}) = \begin{bmatrix} \frac{\partial}{\partial p_1} h_1(\mathrm{p}, \bar{u}) & \dots & \frac{\partial}{\partial p_n} h_1(\mathrm{p}, \bar{u}) \\ \vdots & & \vdots \\ \frac{\partial}{\partial p_1} h_n(\mathrm{p}, \bar{u}) & \dots & \frac{\partial}{\partial p_n} h_n(\mathrm{p}, \bar{u}) \end{bmatrix} \tag{19.7}$$

Call it the *substitution matrix* since it contains all the Hicksian substitution terms.

Properties of the Substitution Matrix

Assume Assumption 1 and Assumption 2, and that compensated demands are continuously differentiable. Then:

1. $\sigma_{ij} = \frac{\partial^2}{\partial p_i \partial p_j} e(\mathrm{p}, \bar{u})$.
2. σ is symmetric.
3. *Adding up*: $\sigma \cdot \mathrm{p} = 0$. That is, for all $i = 1, \dots, n$, $\sum_{j=1}^{n} p_j \frac{\partial}{\partial p_j} h_i(\mathrm{p}, \bar{u}) = 0$.
4. σ is negative semi-definite.[2]

Exercise CD.1

Use these properties to prove that each good has at least one "net substitute," that is, for all i, there exists a j such that $\frac{\partial}{\partial p_j} h_i(\mathrm{p}, \bar{u}) > 0$.

The requirements that consumer demand satisfy homogeneity and budget balancedness, and that the Slutsky matrix be symmetric and negative semi-definite, provide a set of *restrictions* on the parameters in any empirically estimated Marshallian demand system.

[2] An $n \times n$ matrix is *negative semi-definite* if for all vectors $\mathrm{z} \in \mathbb{R}^n$, $\mathrm{z}^T A \mathrm{z} = \sum_{i=1}^{n} \sum_{j=1}^{n} z_j a_{ij} z_i \leq 0$. If the inequality is strict for all non-zero z, A is *negative definite*.

Integrability

If $x^*(p, w)$ is a utility-maximizer's system of demand functions, we may summarize the implications for observable behavior as:

Restrictions Imposed on Demand Behavior

1. Budget Balancedness: $p \cdot x^*(p, w) = w$
2. Negative Semi-definiteness: The associated Slutsky matrix $S(p, w) = w$ must be negative semi-definite
3. Symmetry: $S(p, w)$ must be symmetric

Integrability Theorem

A continuously differentiable function $x^* : \mathbb{R}^{n+1}_{++} \to \mathbb{R}^n_+$ is the demand function generated by some increasing, quasiconcave utility function if (and only if, when utility is continuous, strictly increasing, and strictly quasiconcave) it satisfies budget balancedness, symmetry, and negative semi-definiteness.

Conclusion

If one wishes to estimate a consumer's demand function based on a limited amount of data, and one wishes to impose as a restriction that the demand function be utility-generated, one is now free to specify any functional form for demand as long as it satisfies budget balancedness, symmetry, and negative semi-definiteness.

Figure 19.1: Thesis defense (xkcd.com)

20 Compensating Variation

**Jehle and Philip p28-60*
Rubenstein Lecture 6
Kreps p51-59
MWG 3D-G

Duality

Relationships between programs: Assume Assumption 1 . Given $(\mathbf{p}, w)$, let $\bar{u} = v(\mathbf{p}, w)$. Then

Duality 1

1. $e(\mathbf{p}, v(\mathbf{p}, w)) = w$
2. $v(\mathbf{p}, e(\mathbf{p}, \bar{u})) = \bar{u}$

Further, we have the following relations between the Marshallian and Hicksian demand functions:

Duality 2

1. $\mathbf{x}^*(\mathbf{p}, w) = \mathbf{h}(\mathbf{p}, v(\mathbf{p}, w))$
2. $\mathbf{h}(\mathbf{p}, \bar{u}) = \mathbf{x}^*(\mathbf{p}, e(\mathbf{p}, \bar{u}))$

Suggested Exercise CV.1

Outline the steps of the proof of this proposition of the duality between Marshallian and Hicksian demand functions (it's in your texts if you need help, but try it yourself first). Use an algebraic, not graphical, approach. For example, start with a solution to the utility maximization problem, and show by contradiction that it must also solve the expenditure minimization problem. Then do the same in the other direction.[1]

The Slutsky Equation

Now we can use our knowledge of the substitution *matrix* to make a comprehensive deduction about the effects of *all* price and income changes on the entire system of observable Marshallian demands.

For $i, j = 1, \ldots, n$, the element in the ith row and jth column of the Slutsky Matrix (S) is

$$S_{ij} = \frac{\partial}{\partial p_j} h_i(\mathbf{p}, \bar{u}) = \frac{\partial}{\partial p_j} x_i^*(\mathbf{p}, w) + x_j^*(\mathbf{p}, w) \frac{\partial}{\partial w} x_i^*(\mathbf{p}, w) \tag{20.1}$$

where $\bar{u} = v(\mathbf{p}, w)$.

Exercise CV.2

> Demonstrate that equation (20.1) can be seen by differentiating an identity and applying the envelope theorem.

Exercise CV.3

> Argue that, in contrast to the case of compensated demand, $\frac{\partial}{\partial p_j} x_i^*(\mathbf{p}, w) \neq \frac{\partial}{\partial p_i} x_j^*(\mathbf{p}, w)$.

The Fundamental Equation of Demand Theory

Let $\mathbf{x}^*(\mathbf{p}, w)$ be the consumer's Marshallian demand system. Let u^* be the level of utility the consumer achieves at prices $\mathbf{p}$ and income w. Then,

$$\underbrace{\frac{\partial x_i^*(\mathbf{p}, w)}{\partial p_j}}_{TE} = \underbrace{\frac{\partial h_i(\mathbf{p}, u^*)}{\partial p_j}}_{SE} \underbrace{- x_j^*(\mathbf{p}, w) \frac{\partial x_i^*(\mathbf{p}, w)}{\partial w}}_{IE}, \qquad i, j = 1, \ldots, n. \tag{20.2}$$

Part of this "Fundamental Equation" is fundamentally unobservable, since we cannot observe Hicksian demand curves. Nevertheless, our theory tells us quite a bit about Hicksian demands.

So What?

Before we start the next section, let's remember why we're doing this. The ultimate goal: understand human behavior. Why? One reason is to design and implement policies that maximize social well-being. Other reasons?

Ours is a slightly less lofty goal, although it's a stepping stone to the ultimate goal; we just want to understand consumer demand in the marketplace. In particular, we would like to come up with *testable* hypotheses about consumer behavior. To do this, we have imposed structure on the problem. What comes next is a fundamental component of our goal.

Income and Substitution Effects

Graphs Illustrate the relationship between the Marshallian and Hicksian demand curves by deriving them graphically from indifference curves and budget constraints. Use this to motivate the "Fundamental Equation of Demand Theory." Also use it to illustrate Compensating Variation.

- Derive the Marshallian demand curve for good 1 from utility max by raising and lowering the price of good 1
- This incorporates two distinct types of behavior.
  1. One is re-balancing the basket because the relative prices have changed.
  2. The other is buying less (or more) b/c you're richer in real terms.
- We want to isolate the "re-balancing." Why? Hopefully that will be clear in a minute.
- Think about the same price changes as above, but this time the consumer will be compensated.[2] In other words, draw three budget constraints, one for each price, all tangent to the same indifference curve.
- At the higher price, will compensated demand for good 1 be greater than or less than uncompensated (Marshallian) demand? It will be greater (unless the good is inferior). Graph this point on the same graph as the Marshallian demand curve.
- At the lower price, will compensated demand for good 1 be greater or less than the Marshallian demand? It will be less. Graph this point too.
- Connect these compensated demand points. This is the Hicksian demand curve. The Hicksian demand curve picks up precisely the pure Hicksian substitution effect of an own-price change.

[2] It's more intuitive to think about a price increase followed by "compensation" paid to the consumer so he is no worse off.

- Finish it off by illustrating the substitution and income effects for normal, inferior, and strongly inferior (Giffen) goods.

Compensating Variation

That was fun. But why do we care? Remember our town with the mass transit proposal? Well, consider a resident whose income is w^0, and suppose the initial price of transportation around town is p^0 and it will fall to p^1 as a result of the mass transit system. $v(p^0, w^0)$ is the resident's utility before the price fall, and $v(p^1, w^0)$ is his utility after. The amount of income the consumer is willing to give up for the price decrease will be just enough so that at the lower price and income levels he would be just as well off as at the initial higher price and income levels. Let CV denote the change in the resident's income that would leave him as well off after the price fall as before. We have

$$v(p^1, w^0 + CV) = v(p^0, w^0) \tag{20.3}$$

where $CV \leq 0$ in this scenario.

Exploiting duality, we have

$$e(p^1, v(p^0, w^0)) = e(p^1, v(p^1, w^0 + CV)) = w^0 + CV \tag{20.4}$$

Our first duality result (Duality 1.1) tell us that $w^0 = e(p^0, v(p^0, w^0))$. Substitute this into equation (20.4) and rearrange to get

$$\begin{aligned} CV &= e(p^1, v^0) - e(p^0, v^0) && (20.5)\\ &= \int_{p^0}^{p^1} \frac{\partial e(p, v^0)}{\partial p} dp && (20.6)\\ &= \int_{p^0}^{p^1} h(p, v^0) dp && (20.7) \end{aligned}$$

This is the area "under" the Hicksian demand curve between p^0 and p^1. It is called *Compensating Variation*, and it is a dollar-denominated measure of the welfare impact a price change will have. We got this by going from

$$\max_{\mathbf{x}} u \quad \longrightarrow \quad v(\mathbf{p}, w) \quad \longrightarrow \quad e(\mathbf{p}, \bar{u}) \quad \longrightarrow \quad \mathbf{h}(\mathbf{p}, w). \tag{20.8}$$

Pretty cool, huh?

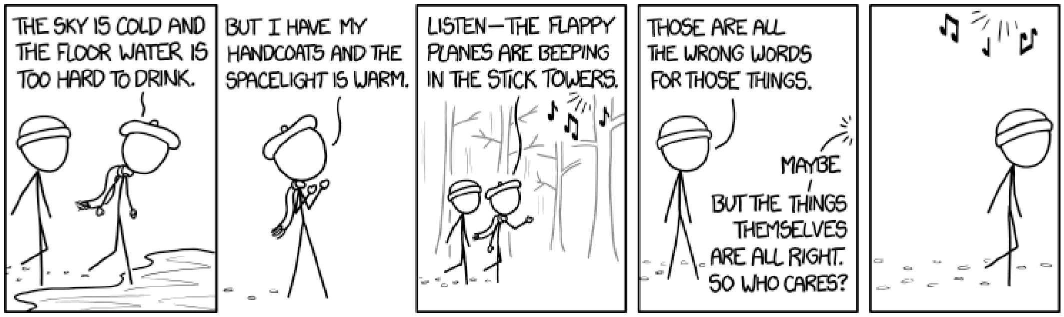

Figure 20.1: Winter (xkcd.com)